

Connecting via Winsock to STN

Welcome to STN International! Enter x:x

LOGINID:SSSPTA1653HXP

PASSWORD:

TERMINAL (ENTER 1, 2, 3, OR ?):2

* * * * * Welcome to STN International * * * * *

NEWS 1 Web Page URLs for STN Seminar Schedule - N. America
NEWS 2 "Ask CAS" for self-help around the clock
NEWS 3 JUL 20 Powerful new interactive analysis and visualization software,
STN AnaVist, now available
NEWS 4 AUG 11 Derwent World Patents Index(R) web-based training during
August
NEWS 5 AUG 11 STN AnaVist workshops to be held in North America
NEWS 6 AUG 30 CA/CAPLUS - Increased access to 19th century research documents
NEWS 7 AUG 30 CASREACT - Enhanced with displayable reaction conditions
NEWS 8 SEP 09 ACD predicted properties enhanced in REGISTRY/ZREGISTRY

NEWS EXPRESS JUNE 13 CURRENT WINDOWS VERSION IS V8.0, CURRENT
MACINTOSH VERSION IS V6.0c(ENG) AND V6.0Jc(JP),
AND CURRENT DISCOVER FILE IS DATED 13 JUNE 2005

NEWS HOURS STN Operating Hours Plus Help Desk Availability
NEWS INTER General Internet Information
NEWS LOGIN Welcome Banner and News Items
NEWS PHONE Direct Dial and Telecommunication Network Access to STN
NEWS WWW CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that
specific topic.

All use of STN is subject to the provisions of the STN Customer
agreement. Please note that this agreement limits use to scientific
research. Use for software development or design or implementation
of commercial gateways or other similar uses is prohibited and may
result in loss of user privileges and other penalties.

* * * * * STN Columbus * * * * *

FILE 'HOME' ENTERED AT 13:07:27 ON 20 SEP 2005

=> file medline, uspatful, dgene, embase, wpids, biosis, biotechds
COST IN U.S. DOLLARS SINCE FILE TOTAL
ENTRY SESSION
FULL ESTIMATED COST 0.21 0.21

FILE 'MEDLINE' ENTERED AT 13:08:15 ON 20 SEP 2005

FILE 'USPATFULL' ENTERED AT 13:08:15 ON 20 SEP 2005
CA INDEXING COPYRIGHT (C) 2005 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'DGENE' ENTERED AT 13:08:15 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

FILE 'EMBASE' ENTERED AT 13:08:15 ON 20 SEP 2005
COPYRIGHT (C) 2005 Elsevier Inc. All rights reserved.

FILE 'WPIDS' ENTERED AT 13:08:15 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

FILE 'BIOSIS' ENTERED AT 13:08:15 ON 20 SEP 2005
Copyright (c) 2005 The Thomson Corporation

FILE 'BIOTECHDS' ENTERED AT 13:08:15 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

=> s (integrating plasmid)
L1 272 (INTEGRATING PLASMID)

=> s candida utilis
L2 5634 CANDIDA UTILIS

=> s l2 and (biotin synthase)
L3 11 L2 AND (BIOTIN SYNTHASE)

=> d l3 ti abs ibib tot

L3 ANSWER 1 OF 11 USPATFULL on STN
TI Methods for monitoring multiple gene expression
AB The present invention relates to methods for monitoring differential expression of a plurality of genes in a first filamentous fungal cell relative to expression of the same genes in one or more second filamentous fungal cells using microarrays containing filamentous fungal expressed sequenced tags. The present invention also relates to filamentous fungal expressed sequenced tags and to computer readable media and substrates containing such expressed sequenced tags for monitoring expression of a plurality of genes in filamentous fungal cells.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2005:140340 USPATFULL
TITLE: Methods for monitoring multiple gene expression
INVENTOR(S): Berka, Randy M., Davis, CA, UNITED STATES
Rey, Michael W., Davis, CA, UNITED STATES
Shuster, Jeffrey R., Davis, CA, UNITED STATES
Kauppinen, Sakari, Smoerum, DENMARK
Clausen, Ib Groth, Hillerod, DENMARK
Olsen, Peter Bjarke, Copenhagen, DENMARK
PATENT ASSIGNEE(S): Novozymes Biotech, Inc., Davis, CA, UNITED STATES (U.S. corporation)
Novozymes A/S, Bagsvaerd, DENMARK (non-U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 6902887	B1	20050607
APPLICATION INFO.:	US 2000-533559		20000322 (9)
RELATED APPLN. INFO.:	Continuation-in-part of Ser. No. US 1999-273623, filed on 22 Mar 1999, ABANDONED		
DOCUMENT TYPE:	Utility		
FILE SEGMENT:	GRANTED		
PRIMARY EXAMINER:	Brusca, John S.		
LEGAL REPRESENTATIVE:	Stames, Robert L.		
NUMBER OF CLAIMS:	8		
EXEMPLARY CLAIM:	1		
NUMBER OF DRAWINGS:	0 Drawing Figure(s); 0 Drawing Page(s)		
LINE COUNT:	18386		

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L3 ANSWER 2 OF 11 USPATFULL on STN

TI Methods for monitoring multiple gene expression

AB The present invention relates to methods for monitoring differential expression of a plurality of genes in a first filamentous fungal cell relative to expression of the same genes in one or more second filamentous fungal cells using microarrays containing filamentous fungal expressed sequenced tags. The present invention also relates to filamentous fungal expressed sequenced tags and to computer readable media and substrates containing such expressed sequenced tags for monitoring expression of a plurality of genes in filamentous fungal cells.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2004:292283 USPATFULL

TITLE: Methods for monitoring multiple gene expression

INVENTOR(S): Berka, Randy M., Davis, CA, UNITED STATES

Rey, Michael W., Davis, CA, UNITED STATES

Shuster, Jeffrey R., Davis, CA, UNITED STATES

Kauppinen, Sakari, Smoerum, DENMARK

Clausen, Ib Groth, Hillerod, DENMARK

Olsen, Peter Bjarke, Copenhagen, DENMARK

PATENT ASSIGNEE(S): Novozymes Biotech, Inc., Davis, CA (U.S. corporation)
Novozymes A/S, Inc., Bagsvaerd, DENMARK (U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2004229367	A1	20041118
APPLICATION INFO.:	US 2003-653047	A1	20030829 (10)
RELATED APPLN. INFO.:	Division of Ser. No. US 2000-533559, filed on 22 Mar 2000, PENDING Continuation-in-part of Ser. No. US 1999-273623, filed on 22 Mar 1999, ABANDONED		
DOCUMENT TYPE:	Utility		
FILE SEGMENT:	APPLICATION		
LEGAL REPRESENTATIVE:	NOVOZYMES BIOTECH, INC., 1445 DREW AVE, DAVIS, CA, 95616		
NUMBER OF CLAIMS:	8		
EXEMPLARY CLAIM:	1		
LINE COUNT:	19697		

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L3 ANSWER 3 OF 11 USPATFULL on STN

TI Yeast with high biotin-productivity and the preparation method thereof

AB The invention discloses a yeast with high biotin-productivity and the preparation method thereof. The yeast is transformed by an integrated plasmid, which includes a **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2003:152918 USPATFULL

TITLE: Yeast with high biotin-productivity and the preparation method thereof

INVENTOR(S): Shiuan, David, Kaohsiung, TAIWAN, PROVINCE OF CHINA

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2003104584	A1	20030605
APPLICATION INFO.:	US 2001-752957	A1	20010102 (9)

NUMBER	DATE
--------	------

PRIORITY INFORMATION: TW 2000-89120972 20001007
 DOCUMENT TYPE: Utility
 FILE SEGMENT: APPLICATION
 LEGAL REPRESENTATIVE: BRUCE LONDA, NORRIS, MCLAUGHLIN & MARCUS, P.A., 220
 EAST 42ND STREET, 30TH FLOOR, NEW YORK, NY, 10017
 NUMBER OF CLAIMS: 16
 EXEMPLARY CLAIM: 1
 NUMBER OF DRAWINGS: 13 Drawing Page(s)
 LINE COUNT: 486
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L3 ANSWER 4 OF 11 USPATFULL on STN

TI Novel Polynucleotides

AB Novel polynucleotides derived from microorganisms belonging to
 coryneform bacteria and fragments thereof, polypeptides encoded by the
 polynucleotides and fragments thereof, polynucleotide arrays comprising
 the polynucleotides and fragments thereof, recording media in which the
 nucleotide sequences of the polynucleotide and fragments thereof have
 been recorded which are readable in a computer, and use of them.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2002:343879 USPATFULL
 TITLE: Novel Polynucleotides
 INVENTOR(S): Nakagawa, Satoshi, Tokyo, JAPAN
 Mizoguchi, Hiroshi, Tokyo, JAPAN
 Ando, Seiko, Tokyo, JAPAN
 Hayashi, Mikiro, Tokyo, JAPAN
 Ochiai, Keiko, Tokyo, JAPAN
 Yokoi, Haruhiko, Tokyo, JAPAN
 Tateishi, Naoko, Tokyo, JAPAN
 Senoh, Akihiro, Tokyo, JAPAN
 Ikeda, Masato, Tokyo, JAPAN
 Ozaki, Akio, Hofu-shi, JAPAN

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2002197605	A1	20021226
APPLICATION INFO.:	US 2000-738626	A1	20001218 (9)

	NUMBER	DATE
PRIORITY INFORMATION:	JP 1999-377484	19991216
	JP 2000-159162	20000407
	JP 2000-280988	20000803
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	NIXON & VANDERHYE P.C., 8th Floor, 1100 North Glebe Road, Arlington, VA, 22201	
NUMBER OF CLAIMS:	68	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	4 Drawing Page(s)	
LINE COUNT:	13673	
CAS INDEXING IS AVAILABLE FOR THIS PATENT.		

L3 ANSWER 5 OF 11 DGENE COPYRIGHT 2005 The Thomson Corp on STN

TI New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.

AN ADA09589 DNA DGENE

AB The present invention relates to a yeast with high biotin-productivity
 and a method for its preparation. The yeast is transformed by an
 integrated plasmid, which includes the edible yeast (*Candida
 utilis*) biotin synthase gene, an assistant

DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents a PCR primer used in the examples of the present invention.

ACCESSION NUMBER: ADA09589 DNA DGENE
TITLE: New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.
INVENTOR: Shiuan D
PATENT ASSIGNEE: (SHIU-I)SHIUAN D.
PATENT INFO: US 2003104584 A1 20030605 20
APPLICATION INFO: US 2001-752957 20010102
PRIORITY INFO: TW 2000-120972 20001007
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2003-567956 [53]
DESCRIPTION: PCR primer #1 for *S. cerevisiae* BIO2 gene.

L3 ANSWER 6 OF 11 DGENE COPYRIGHT 2005 The Thomson Corp on STN
TI New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.
AN ADA09591 DNA DGENE
AB The present invention relates to a yeast with high biotin-productivity and a method for its preparation. The yeast is transformed by an integrated plasmid, which includes the edible yeast (*Candida utilis*) **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents a degenerate PCR primer used in the examples of the present invention.

ACCESSION NUMBER: ADA09591 DNA DGENE
TITLE: New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.
INVENTOR: Shiuan D
PATENT ASSIGNEE: (SHIU-I)SHIUAN D.
PATENT INFO: US 2003104584 A1 20030605 20
APPLICATION INFO: US 2001-752957 20010102
PRIORITY INFO: TW 2000-120972 20001007
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2003-567956 [53]
DESCRIPTION: Degenerate PCR primer #1 for *C. utilis* BIO2 gene.

L3 ANSWER 7 OF 11 DGENE COPYRIGHT 2005 The Thomson Corp on STN
TI New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.
AN ADA09590 DNA DGENE
AB The present invention relates to a yeast with high biotin-productivity and a method for its preparation. The yeast is transformed by an integrated plasmid, which includes the edible yeast (*Candida utilis*) **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents a PCR primer used in the examples of the present invention.

ACCESSION NUMBER: ADA09590 DNA DGENE
TITLE: New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.

INVENTOR: Shiuan D
PATENT ASSIGNEE: (SHIU-I)SHIUAN D.
PATENT INFO: US 2003104584 A1 20030605 20
APPLICATION INFO: US 2001-752957 20010102
PRIORITY INFO: TW 2000-120972 20001007
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2003-567956 [53]
DESCRIPTION: PCR primer #2 for S. cerevisiae BIO2 gene.

L3 ANSWER 8 OF 11 DGENE COPYRIGHT 2005 The Thomson Corp on STN
TI New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.
AN ADA09592 DNA DGENE
AB The present invention relates to a yeast with high biotin-productivity and a method for its preparation. The yeast is transformed by an integrated plasmid, which includes the edible yeast (**Candida utilis**) **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents a degenerate PCR primer used in the examples of the present invention.

ACCESSION NUMBER: ADA09592 DNA DGENE
TITLE: New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.

INVENTOR: Shiuan D
PATENT ASSIGNEE: (SHIU-I)SHIUAN D.
PATENT INFO: US 2003104584 A1 20030605 20
APPLICATION INFO: US 2001-752957 20010102
PRIORITY INFO: TW 2000-120972 20001007
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2003-567956 [53]
DESCRIPTION: Degenerate PCR primer #2 for C. utilis BIO2 gene.

L3 ANSWER 9 OF 11 DGENE COPYRIGHT 2005 The Thomson Corp on STN
TI New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.
AN ADA09588 DNA DGENE
AB The present invention relates to a yeast with high biotin-productivity and a method for its preparation. The yeast is transformed by an integrated plasmid, which includes the edible yeast (**Candida utilis**) **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents C. utilis **biotin synthase** gene.

ACCESSION NUMBER: ADA09588 DNA DGENE
TITLE: New integrated plasmid, useful for preparing yeast with high biotin- productivity used as feed additives, food additives or cosmetics.

INVENTOR: Shiuan D
PATENT ASSIGNEE: (SHIU-I)SHIUAN D.
PATENT INFO: US 2003104584 A1 20030605 20
APPLICATION INFO: US 2001-752957 20010102
PRIORITY INFO: TW 2000-120972 20001007
DOCUMENT TYPE: Patent
LANGUAGE: English
OTHER SOURCE: 2003-567956 [53]
DESCRIPTION: **Candida utilis biotin**

synthase gene.

L3 ANSWER 10 OF 11 WPIDS COPYRIGHT 2005 THE THOMSON CORP on STN
TI New integrated plasmid, useful for preparing yeast with high
biotin-productivity used as feed additives, food additives or cosmetics.
AN 2003-567956 [53] WPIDS
AB US2003104584 A UPAB: 20030820

NOVELTY - An integrated plasmid comprising a **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for the following:

(1) a method for preparing a yeast with high biotin-productivity comprising constructing the integrated plasmid cited above, linearizing the integrated plasmid, transforming the linearized integrated plasmid into a yeast, and recombining the **biotin synthase** gene with the yeast genome; and

(2) a method for producing biotin comprising providing the yeast with high biotin-productivity, culturing the yeast in a nutrient medium, and recovering biotin from the culture broth.

USE - The integrated plasmid is useful for preparing yeast with high biotin-productivity. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics (all claimed).

Dwg.0/13

ACCESSION NUMBER: 2003-567956 [53] WPIDS
DOC. NO. CPI: C2003-153236
TITLE: New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.
DERWENT CLASS: C06 D13 D16 D21
INVENTOR(S): SHIUAN, D
PATENT ASSIGNEE(S): (NASC-N) NAT SCI COUNCIL; (SHIU-I) SHIUAN D
COUNTRY COUNT: 2
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
US 2003104584	A1	20030605	(200353)*		20
TW 220675	B1	20040901	(200522)		

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
US 2003104584	A1	US 2001-752957	20010102
TW 220675	B1	TW 2000-120972	20001007

PRIORITY APPLN. INFO: TW 2000-120972 20001007

L3 ANSWER 11 OF 11 BIOTECHDS COPYRIGHT 2005 THE THOMSON CORP. on STN
TI New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics; vector-mediated gene transfer and expression in host cell for strain improvement and feed additive, food additive or cosmetic manufacture
AN 2003-20688 BIOTECHDS
AB DERWENT ABSTRACT:
NOVELTY - An integrated plasmid comprising a **biotin synthase** gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are included for the following: (1) a method for preparing a yeast with high biotin-productivity comprising constructing the integrated plasmid cited above, linearizing the integrated plasmid, transforming the linearized integrated plasmid into a yeast, and recombining the **biotin synthase** gene with the yeast genome; and (2) a method for producing biotin comprising providing the yeast with high biotin-productivity, culturing the yeast in a nutrient medium, and recovering biotin from the culture broth.

BIOTECHNOLOGY - Preferred Plasmid: The **biotin synthase** gene of the integrated plasmid is derived from *Saccharomyces cerevisiae* or *Candida utilis*. The **biotin synthase** gene of *C. utilis* comprises a fully defined sequence of 1188 base pairs (bp) given in the specification. The assistant DNA sequence is a *C. utilis* fragment consisting of *NsiI*-*BamHI* 18s rDNA, *URA3* DNA, or *HIS3* DNA. The selection marker is preferably a cycloheximide-resistant gene. The promoter sequence is pL41 promoter of *C. utilis* or pADH1 promoter of *S. cerevisiae*. The integrated plasmid is: (A) pMCC21 comprising pL41 for the expression of *C. utilis* BIO2 gene encoding for 204 amino acids, mutated L41 gene used as cycloheximide-resistant marker, and *C. utilis* 18S rDNA; (B) pMCC31S or pMCC38S comprising pL41 for the expression of *C. utilis* BIO2 gene encoding for 233 amino acids, mutated L41 gene used as cycloheximide-resistant marker, and *C. utilis* 18S rDNA; (C) pMCC32H or pMCC36H comprising *C. utilis* pL41, *C. utilis* BIO2 gene, mutated L41 gene used as cycloheximide-resistant marker, and *C. utilis* *HIS3* gene fragment; or (D) pMCC33U or pMCC35U comprising *C. utilis* pL41, *C. utilis* BIO2 gene, mutated L41 gene used as cycloheximide-resistant marker, and *C. utilis* *URA3* gene fragment.

USE - The integrated plasmid is useful for preparing yeast with high biotin-productivity. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics (all claimed).

EXAMPLE - Polymerase chain reaction (PCR) was carried out using the *Candida utilis* BIO2 gene as a template. The product was digested with *NdeI* and *SalI*, and ligated into pMC5 to obtain pMCC11. The *NdeI*-*SalI* fragment of pMCC11 in which the *NdeI* nick was end-repaired into a blunt end by Klenow DNA polymerase (i.e. pL41-BIO2 fragment) was ligated into pMC9 to obtain pMCC15. To fill the lost L41 fragment, the *BamHI*-*EcoRI* fragment of pMC9 (i.e. L41 gene fragment) was end-repaired into a blunt end by Klenow DNA polymerase and then integrated into the *SalI* site in pMCC15 to obtain pMCC21. The plasmid pMCC21 comprises *C. utilis* pL41 for the expression of the *C. utilis* BIO2 gene, mutated L41 gene used as cycloheximide-resistant marker, and *C. utilis* 18S rDNA. (20 pages)

ACCESSION NUMBER: 2003-20688 BIOTECHDS

TITLE: New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics;
vector-mediated gene transfer and expression in host cell for strain improvement and feed additive, food additive or cosmetic manufacture

AUTHOR: SHIUAN D

PATENT ASSIGNEE: SHIUAN D

PATENT INFO: US 2003104584 5 Jun 2003

APPLICATION INFO: US 2001-752957 2 Jan 2001

PRIORITY INFO: TW 2000-120972 7 Oct 2000; TW 2000-120972 7 Oct 2000

DOCUMENT TYPE: Patent

LANGUAGE: English

OTHER SOURCE: WPI: 2003-567956 [53]

=> s yeast preparation

L4 310 YEAST PREPARATION

=> s 14 and (biotin production)
L5 0 L4 AND (BIOTIN PRODUCTION)

=> s 14 and (biotin)
L6 8 L4 AND (BIOTIN)

=> d 16 ti abs ibib tot

L6 ANSWER 1 OF 8 USPATFULL on STN

TI Whole cell engineering using real-time metabolic flux analysis
AB The invention provides methods for whole cell engineering of new and modified phenotypes by using "on-line" or "real-time" metabolic flux analysis. The invention provides a method for whole cell engineering of new or modified phenotypes by using real-time metabolic flux analysis by making a modified cell by modifying the genetic composition of a cell and culturing the modified cell to generate a plurality of modified cells and measuring at least one metabolic parameter of the cell by monitoring the cell culture of in real time. The invention also provides articles comprising machine-readable medium including machine-executable instructions and systems, e.g., computer systems, to practice the methods of the invention.

ACCESSION NUMBER: 2005:233489 USPATFULL
TITLE: Whole cell engineering using real-time metabolic flux analysis
INVENTOR(S): Short, Jay M., Rancho Sante Fe, CA, UNITED STATES
Latterich, Martin, Montreal, CANADA
Levin, Michael, San Diego, CA, UNITED STATES
Wei, Jing, San Diego, CA, UNITED STATES
Fu, Pengcheng, Honolulu, HI, UNITED STATES

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2005202426	A1	20050915
APPLICATION INFO.:	US 2003-491358	A1	20021001 (10)
	WO 2002-US31380		20021001
			20050223 PCT 371 date

	NUMBER	DATE
PRIORITY INFORMATION:	US 2003-326653P	20011001 (60)
	US 2003-326654P	20011001 (60)
	US 2003-326655P	20011001 (60)
	US 2003-337526P	20011109 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	DIVERSA C/O MOFO S.D., 3811 VALLEY CENTER DRIVE, SUITE 500, SAN DIEGO, CA, 92130, US	
NUMBER OF CLAIMS:	80	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	40 Drawing Page(s)	
LINE COUNT:	4859	

L6 ANSWER 2 OF 8 USPATFULL on STN

TI Screening method for genes of brewing yeast
AB The present invention provides (A) a method for the selection of genes participating in the desired brewing character, which comprises preparing a database compiling the data of the whole genome sequence of industrial yeast, particularly a brewing yeast used for alcoholic beverages; selecting gene participating in a desired brewing character that the brewing yeast specifically possesses; and carrying out

functional analysis of the gene by disruption or overexpression; (B) a DNA array in which oligonucleotide(s) selected based on the data base compiling the data of the whole genome sequences of an industrial yeast, (C) a breeding method for constructing improved cultures achieving the desired brewing character, (D) a method for producing an alcohol or an alcoholic beverage in which productivity and quality are improved using the yeast, (E) a gene which is specific to the improved brewing yeast, and (F) a peptide encoded by the gene.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2004:334780 USPATFULL
TITLE: Screening method for genes of brewing yeast
INVENTOR(S): Nakao, Yoshihiro, Kyoto-shi, JAPAN
Nakamura, Norihisa, Kyoto-shi, JAPAN
Kodama, Yukiko, Osaka, JAPAN
Fujimura, Tomoko, Osaka, JAPAN
Ashikari, Toshihiko, Osaka, JAPAN
PATENT ASSIGNEE(S): SUNTORY LIMITED, OSAKA, JAPAN (non-U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2004265862	A1	20041230
APPLICATION INFO.:	US 2004-791791	A1	20040304 (10)

	NUMBER	DATE
PRIORITY INFORMATION:	JP 2003-57677	20030304
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	BURNS DOANE SWECKER & MATHIS L L P, POST OFFICE BOX 1404, ALEXANDRIA, VA, 22313-1404	
NUMBER OF CLAIMS:	46	
EXEMPLARY CLAIM:	1	
NUMBER OF DRAWINGS:	12 Drawing Page(s)	
LINE COUNT:	3038	

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L6 ANSWER 3 OF 8 USPATFULL on STN

TI Whole cell engineering using real-time metabolic flux analysis
AB The invention provides methods for whole cell engineering of new and modified phenotypes by using "on-line" or "real-time" metabolic flux analysis. The invention provides a method for whole cell engineering of new or modified phenotypes by using real-time metabolic flux analysis by making a modified cell by modifying the genetic composition of a cell and culturing the modified cell to generate a plurality of modified cells and measuring at least one metabolic parameter of the cell by monitoring the cell culture of in real time. The invention also provides articles comprising machine-readable medium including machine-executable instructions and systems, e.g., computer systems, to practice the methods of the invention.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2004:44982 USPATFULL
TITLE: Whole cell engineering using real-time metabolic flux analysis
INVENTOR(S): Fu, Pengcheng, Honolulu, HI, UNITED STATES
Latterich, Martin, San Diego, CA, UNITED STATES
Levin, Michael, San Diego, CA, UNITED STATES
Wei, Jing, San Diego, CA, UNITED STATES
Short, Jay M., Rancho Santa Fe, CA, UNITED STATES
PATENT ASSIGNEE(S): Diversa Corporation, San Diego, CA (U.S. corporation)

NUMBER	KIND	DATE
--------	------	------

PATENT INFORMATION:	US 2004033975	A1	20040219
APPLICATION INFO.:	US 2002-264989	A1	20021001 (10)

	NUMBER	DATE
PRIORITY INFORMATION:	US 2001-326655P	20011001 (60)
	US 2001-326654P	20011001 (60)
	US 2001-326653P	20011001 (60)
	US 2001-337526P	20011109 (60)

DOCUMENT TYPE: Utility
FILE SEGMENT: APPLICATION
LEGAL REPRESENTATIVE: GREGORY P. EINHORN, Fish & Richardson P.C., Suite 500,
4350 La Jolla Village Drive, San Diego, CA, 92122
NUMBER OF CLAIMS: 228
EXEMPLARY CLAIM: 1
NUMBER OF DRAWINGS: 40 Drawing Page(s)
LINE COUNT: 5427
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L6 ANSWER 4 OF 8 USPATFULL on STN
TI Baker's yeast and a method producing it
AB A process is disclosed for the production of baker's yeast which
comprises the fermentation of a baker's yeast strain using a
non-molasses carbon source until a fermentation broth having at least
10%, preferably at least 13%, more preferably at least 16% of dry solids
content is formed, which can be used directly as a cream yeast without
concentration.

ACCESSION NUMBER: 1999:72304 USPATFULL
TITLE: Baker's yeast and a method producing it
INVENTOR(S): Plomp, Pieter Jan Arnoldus Maria, Ex Delft, Netherlands
PATENT ASSIGNEE(S): Gist-brocades, B.V., Netherlands (non-U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 5916609		19990629
APPLICATION INFO.:	US 1997-901312		19970728 (8)

	NUMBER	DATE
PRIORITY INFORMATION:	EP 1996-202119	19960727
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	Granted	
PRIMARY EXAMINER:	Sayala, Chhaya D.	
LEGAL REPRESENTATIVE:	Morrison & Foerster, LLP	
NUMBER OF CLAIMS:	21	
EXEMPLARY CLAIM:	1	
LINE COUNT:	537	

L6 ANSWER 5 OF 8 USPATFULL on STN
TI Production of distillers yeast
AB Distillers yeast is produced by a process which results in more rapid
yeast growth, higher cell concentrations and greatly extended yeast
viability. The process involves propagating yeast on a culture medium of
cooked cereal grain mash which has been subjected to enzymatic action of
glucamylase obtained from fungi of the genus Aspergillus. After
propagation, as a result of substantially longer viability, the yeast
can be stored in the culture medium for up to 3 days or longer with
portions being withdrawn from time to time for inoculating distillery
grain mashes.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 75:10033 USPATFULL
TITLE: Production of distillers yeast
INVENTOR(S): Van Lanen, deceased, James M., late of Peoria, IL,
United States
Smith, Merritt B., Peoria, IL, United States
Maisch, Weldon F., Peoria, IL, United States
PATENT ASSIGNEE(S): Hiram Walker & Sons, Inc., Peoria, IL, United States
(U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 3868307		19750225
APPLICATION INFO.:	US 1973-388359		19730815 (5)
RELATED APPLN. INFO.:	Continuation-in-part of Ser. No. US 1971-189678, filed on 15 Oct 1971, now abandoned		
DOCUMENT TYPE:	Utility		
FILE SEGMENT:	Granted		
PRIMARY EXAMINER:	Naff, David M.		
LEGAL REPRESENTATIVE:	Kirkland and Ellis		
NUMBER OF CLAIMS:	6		
EXEMPLARY CLAIM:	1		
NUMBER OF DRAWINGS:	2 Drawing Figure(s); 2 Drawing Page(s)		
LINE COUNT:	633		
CAS INDEXING IS AVAILABLE FOR THIS PATENT.			

L6 ANSWER 6 OF 8 WPIDS COPYRIGHT 2005 THE THOMSON CORP on STN
TI Bread-baking yeast preparation - involves adding aryl-thioacetic acid derivs.
to des thio-**biotin** used in molasses-based nutrient medium for
yeast cultivation.

AN 1988-159690 [23] WPIDS

AB SU 1350172 A UPAB: 19930923

The bread-baking yeasts are obtd. by multistage cultivation under aeration conditions on a nutrient medium containing molasses, mineral salts and desthiobiotin. During the stages of production of yeast inoculant and commercial yeasts, derivs. of arylthio (sulphonyl) acetic acid are added in amount (1-10000)x10 power minus 10% jointly with desthiobiotin.

The specified derivs. are 2-hydroxyalkyl-ammonium salts of formula (I), where (a) R=H, R2=4CH3, n=2 and X=N+H3CH(CH2OH)CH(OH)C6H4NO2-4; (b) R1=H, R2=4Cl, n=0 and X=N+H(CH3)2CH2CH2OH; or (c) R1=H, R2=4Cl, n=2, and X=N+H(CH2CH2OH)3.

(I) contain metabolically active sulphur which can be incorporated into desthiobiotin and form **biotin**. (I) also act as antioxidants by lowering the intensity of peroxide oxidation of lipids of biological membranes.

USE/ADVANTAGE - In the microbiological industry. Increased biomass yield is obtd. Bul.41/7.11.87
0/0

ACCESSION NUMBER: 1988-159690 [23] WPIDS

DOC. NO. CPI: C1988-071563

TITLE: Bread-baking yeast preparation - involves adding
aryl-thioacetic acid derivs. to des thio-**biotin**
used in molasses-based nutrient medium for yeast
cultivation.

DERWENT CLASS: D11 D16 E19

INVENTOR(S): GUSEVA, S A; LEVKOVSKAY, G G; MIRSKOVA, A N

PATENT ASSIGNEE(S): (ASII) AS SIBE IRKUT ORG CHEM; (BREA) LENG D BREAD BAKING
IND

COUNTRY COUNT: 1

PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
SU 1350172	A	19871107	(198823)*		3

APPLICATION DETAILS:

PATENT NO	KIND	APPLICATION	DATE
SU 1350172	A	SU 1985-3960649	19851005

PRIORITY APPLN. INFO: SU 1985-3960649 19851005

L6 ANSWER 7 OF 8 WPIDS COPYRIGHT 2005 THE THOMSON CORP on STN
TI Yeast prepared by culturing Hansenula ofunaensis - in medium containing methanol, ethanol and/or glucose as carbon source.
AN 1978-02953A [02] WPIDS
AB JP 52139773 A UPAB: 19930901
The method is characterised by culturing a novel yeast stock Hansenula of unaensis (FERM-P 3494). The yeast stock Hansenula of unaensis can utilise methanol, ethanol and glucose, which are favourable as the starting material for preparing yeast for food use.

The culture is practiced aerobically at 10-41 pref. 30-39 degrees C in the medium containing methanol, ethanol and/or glucose as carbon source at pH 2-9 pref. 3-7 for up to 6 days. As the growth promoter vitamins such as thiamine, **biotin**, etc. and amino acid such as methionine, cysteine, etc. can be applied. After the culture yeast body is separated washed and dried.

ACCESSION NUMBER: 1978-02953A [02] WPIDS
TITLE: Yeast prepared by culturing Hansenula ofunaensis - in medium containing methanol, ethanol and/or glucose as carbon source.
DERWENT CLASS: C03 D13 D16
PATENT ASSIGNEE(S): (MITK) MITSUI TOATSU CHEM INC
COUNTRY COUNT: 1
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
JP 52139773	A	19771121	(197802)*		
JP 56012109	B	19810318	(198115)		

PRIORITY APPLN. INFO: JP 1976-44940 19760422

L6 ANSWER 8 OF 8 WPIDS COPYRIGHT 2005 THE THOMSON CORP on STN
TI Yeast preparation using Torulopsis methanothermophilus - in culture medium containing methanol as carbon source and **biotin**.
AN 1977-19283Y [11] WPIDS
AB JP 52015882 A UPAB: 19930901

Method is effected by inoculating a stock belonging to Torulopsis methanothermophilus which can utilize methanol, into a culture medium containing methanol as carbon source and **biotin**, culturing it for multiplication and separating the yeast body.

The stock can multiply at high temperature and low pH and is considered a novel stock. The stock is trusted to Institute of Microbial Industry with the Trust Number 3158. As teh carbon source crude methanol which contains small amount of lower aldehydes, lower alcohols, etc. can be used. With methanol concentrate >4% the growth of the yest is inhibited, so methanol is added to the culture medium little by little.

ACCESSION NUMBER: 1977-19283Y [11] WPIDS
TITLE: Yeast preparation using Torulopsis methanothermophilus - in culture medium containing methanol as carbon source and **biotin**.
DERWENT CLASS: D16

PATENT ASSIGNEE(S): (MITP) MITSUBISHI PETROCHEMICAL CO LTD
COUNTRY COUNT: 1
PATENT INFORMATION:

PATENT NO	KIND	DATE	WEEK	LA	PG
JP 52015882	A	19770205	(197711)*		

PRIORITY APPLN. INFO: JP 1975-91807 19750728

=> d his

(FILE 'HOME' ENTERED AT 13:07:27 ON 20 SEP 2005)

FILE 'MEDLINE, USPATFULL, DGENE, EMBASE, WPIDS, BIOSIS, BIOTECHDS'
ENTERED AT 13:08:15 ON 20 SEP 2005

L1 272 S (INTEGRATING PLASMID)
L2 5634 S CANDIDA UTILIS
L3 11 S L2 AND (BIOTIN SYNTHASE)
L4 310 S YEAST PREPARATION
L5 0 S L4 AND (BIOTIN PRODUCTION)
L6 8 S L4 AND (BIOTIN)

=> s l1 and l3

L7 0 L1 AND L3

=> s l2 and l1

L8 2 L2 AND L1

=> d l8 ti abs ibib tot

L8 ANSWER 1 OF 2 USPATFULL on STN

TI Production of farnesol and geranylgeraniol

AB The invention provides a biological method of producing farnesol or geranylgeraniol.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2004:144592 USPATFULL

TITLE: Production of farnesol and geranylgeraniol

INVENTOR(S): Millis, James R., Kohler, WI, UNITED STATES

Maurina-Brunker, Julie, Appleton, WI, UNITED STATES

McMullin, Thomas W., Manitowoc, WI, UNITED STATES

PATENT ASSIGNEE(S): Arkion Life Sciences LLC. (U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2004110257	A1	20040610
APPLICATION INFO.:	US 2003-719196	A1	20031120 (10)
RELATED APPLN. INFO.:	Continuation of Ser. No. US 2001-909558, filed on 20 Jul 2001, GRANTED, Pat. No. US 6689593 Continuation of Ser. No. US 1999-350275, filed on 6 Jul 1999, GRANTED, Pat. No. US 6531303		

	NUMBER	DATE
PRIORITY INFORMATION:	US 1998-91964P	19980706 (60)
DOCUMENT TYPE:	Utility	
FILE SEGMENT:	APPLICATION	
LEGAL REPRESENTATIVE:	SHERIDAN ROSS PC, 1560 BROADWAY, SUITE 1200, DENVER, CO, 80202	

NUMBER OF CLAIMS: 17
EXEMPLARY CLAIM: 1
NUMBER OF DRAWINGS: 11 Drawing Page(s)
LINE COUNT: 3450
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L8 ANSWER 2 OF 2 USPATFULL on STN
TI Manufacture of five-carbon sugars and sugar alcohols
AB The invention relates to the methods of manufacturing five-carbon sugars and sugar alcohols as well as other compounds derived from pentose-phosphate pathway from readily available substrates such a hexoses using metabolically engineered microbial hosts.

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ACCESSION NUMBER: 2003:99684 USPATFULL
TITLE: Manufacture of five-carbon sugars and sugar alcohols
INVENTOR(S): Miasnikov, Andrei, Kantvik, FINLAND
Ojamo, Heikki, Kirkkonummi, FINLAND
Povelainen, Mira, Espoo, FINLAND
Gros, Hakan, Kantvik, FINLAND
Toivari, Mervi, Espoo, FINLAND
Richard, Peter, Helsinki, FINLAND
Ruuhonen, Laura, Helsinki, FINLAND
Koivuranta, Kari, Helsinki, FINLAND
Londesborough, John, Helsinki, FINLAND
Aristidou, Aristos, Espoo, FINLAND
Penttila, Merja, Helsinki, FINLAND
Plazagnet-Menut, Claire, Paris, FRANCE
Deutscher, Josef, Fontenay Le Fleury, FRANCE
PATENT ASSIGNEE(S): Xyrofin Oy (non-U.S. corporation)

	NUMBER	KIND	DATE
PATENT INFORMATION:	US 2003068791	A1	20030410
APPLICATION INFO.:	US 2001-908744	A1	20010720 (9)
RELATED APPLN. INFO.:	Continuation-in-part of Ser. No. US 2000-488581, filed on 21 Jan 2000, ABANDONED Continuation-in-part of Ser. No. US 1997-790585, filed on 29 Jan 1997, PENDING Continuation of Ser. No. US 1995-368395, filed on 3 Jan 1995, GRANTED, Pat. No. US 5631150 Continuation of Ser. No. US 1993-110672, filed on 24 Aug 1993, ABANDONED Continuation-in-part of Ser. No. US 1992-973325, filed on 5 Nov 1992, ABANDONED Continuation-in-part of Ser. No. WO 2001-FI51, filed on 22 Jan 2001, UNKNOWN		
DOCUMENT TYPE:	Utility		
FILE SEGMENT:	APPLICATION		
LEGAL REPRESENTATIVE:	STERNE, KESSLER, GOLDSTEIN & FOX PLLC, 1100 NEW YORK AVENUE, N.W., SUITE 600, WASHINGTON, DC, 20005-3934		
NUMBER OF CLAIMS:	166		
EXEMPLARY CLAIM:	1		
NUMBER OF DRAWINGS:	25 Drawing Page(s)		
LINE COUNT:	5422		

CAS INDEXING IS AVAILABLE FOR THIS PATENT.

=> file scisearch, biosis, medline, uspatful, wpids, biotechds, embase, dgene
COST IN U.S. DOLLARS
FULL ESTIMATED COST

SINCE FILE	TOTAL
ENTRY	SESSION
86.18	86.39

FILE 'SCISEARCH' ENTERED AT 13:16:26 ON 20 SEP 2005
Copyright (c) 2005 The Thomson Corporation

FILE 'BIOSIS' ENTERED AT 13:16:26 ON 20 SEP 2005
Copyright (c) 2005 The Thomson Corporation

FILE 'MEDLINE' ENTERED AT 13:16:26 ON 20 SEP 2005

FILE 'USPATFULL' ENTERED AT 13:16:26 ON 20 SEP 2005
CA INDEXING COPYRIGHT (C) 2005 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'WPIDS' ENTERED AT 13:16:26 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

FILE 'BIOTECHDS' ENTERED AT 13:16:26 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

FILE 'EMBASE' ENTERED AT 13:16:26 ON 20 SEP 2005
COPYRIGHT (C) 2005 Elsevier Inc. All rights reserved.

FILE 'DGENE' ENTERED AT 13:16:26 ON 20 SEP 2005
COPYRIGHT (C) 2005 THE THOMSON CORPORATION

=> e clemen, r/au

E1	4	CLEMEN U/AU
E2	1	CLEMEN Y/AU
E3	0 -->	CLEMEN, R/AU
E4	1	CLEMENA G/AU
E5	8	CLEMENA G G/AU
E6	8	CLEMENA W/AU
E7	1	CLEMENA WENDY/AU
E8	1	CLEMENACON G/AU
E9	20	CLEMENCE A/AU
E10	5	CLEMENCE A E/AU
E11	13	CLEMENCE A J/AU
E12	4	CLEMENCE B/AU

=> e gerwig, g/au

E1	2	GERWIG W H/AU
E2	55	GERWIG W H JR/AU
E3	0 -->	GERWIG, G/AU
E4	34	GERWIN B/AU
E5	5	GERWIN B D/AU
E6	310	GERWIN B I/AU
E7	1	GERWIN B L/AU
E8	1	GERWIN BIRGIT/AU
E9	1	GERWIN BRENDA/AU
E10	18	GERWIN BRENDA I/AU
E11	19	GERWIN C/AU
E12	19	GERWIN C M/AU

=> e haess, j/au

E1	1	HAESS K/AU
E2	3	HAESS M/AU
E3	0 -->	HAESS, J/AU
E4	3	HAESSE W/AU
E5	1	HAESSEL DE MENENDEZ G/AU
E6	1	HAESSEL DE MENENDEZ G G/AU
E7	1	HAESSEL DE MENENDEZ GABRIELA G/AU
E8	1	HAESSEL U/AU
E9	7	HAESSEL W/AU
E10	1	HAESSELBA U/AU
E11	1	HAESSELBARTH/AU
E12	1	HAESSELBARTH A/AU

=> e fleischer, b/au

E1	15	FLEISCHER WOLFGANG/AU
E2	2	FLEISCHER Z/AU
E3	0 -->	FLEISCHER, B/AU
E4	1	FLEISCHERDJOLETO O/AU
E5	1	FLEISCHERE ARTHUR C/AU
E6	1	FLEISCHERF/AU
E7	1	FLEISCHERH R/AU
E8	2	FLEISCHERKURTZ G/AU
E9	2	FLEISCHERLAMBROPOULOS E/AU
E10	6	FLEISCHERLAMBROPOULOS H/AU
E11	1	FLEISCHERN R V/AU
E12	1	FLEISCHEROVA J/AU

=> e schwarz, e/au

E1	4	SCHWARZ ZS/AU
E2	1	SCHWARZ ZVI/AU
E3	0 -->	SCHWARZ, E/AU
E4	2	SCHWARZA HG/AU
E5	1	SCHWARZA M J/AU
E6	1	SCHWARZABAUER J E/AU
E7	2	SCHWARZACH R/AU
E8	6	SCHWARZACH T/AU
E9	1	SCHWARZACH T T/AU
E10	2	SCHWARZACHER G/AU
E11	1	SCHWARZACHER G M/AU
E12	1	SCHWARZACHER GERTRUDE MARIA/AU

=> e sigal, l/au

E1	97	SIGAL ZAFRANI B/AU
E2	29	SIGAL ZAFRANI BRIGITTE/AU
E3	0 -->	SIGAL, L/AU
E4	11	SIGALA B/AU
E5	8	SIGALA BARBARA/AU
E6	18	SIGALA C/AU
E7	2	SIGALA DE C C/AU
E8	1	SIGALA DE CAMPOS C/AU
E9	79	SIGALA F/AU
E10	2	SIGALA FILIS/AU
E11	13	SIGALA FRAGISKA/AU
E12	1	SIGALA FRANCHESKA/AU

Refine Search

Search Results -

Terms	Documents
L7 and L1	1

Database:

US Pre-Grant Publication Full-Text Database
US Patents Full-Text Database
US OCR Full-Text Database
EPO Abstracts Database
JPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Search:

L8

Refine Search

Recall Text

Clear

Interrupt

Search History

DATE: Tuesday, September 20, 2005 [Printable Copy](#) [Create Case](#)

Set Name Query

side by side

Hit Count Set Name

result set

DB=PGPB,USPT; PLUR=YES; OP=OR

<u>L8</u>	L7 and l1	1	<u>L8</u>
<u>L7</u>	L6 and (saccharomyces cerevisae)	23150	<u>L7</u>
<u>L6</u>	L5 and (cycloheximide-resistance gene)	73128	<u>L6</u>
<u>L5</u>	L4 and (yeast preparation)	76057	<u>L5</u>
<u>L4</u>	L2 and integrating plasmid	82646	<u>L4</u>
<u>L3</u>	L2 and l1	3	<u>L3</u>
<u>L2</u>	candida utilis biotin synthase	79111	<u>L2</u>
<u>L1</u>	shiuan.in.	76	<u>L1</u>

END OF SEARCH HISTORY

Hit List

[Clear](#)[Generate Collection](#)[Print](#)[Fwd Refs](#)[Bkwd Refs](#)[Generate OACS](#)

Search Results - Record(s) 1 through 1 of 1 returned.

☐ 1. Document ID: US 20030104584 A1

Using default format because multiple data bases are involved.

L8: Entry 1 of 1

File: PGPB

Jun 5, 2003

PGPUB-DOCUMENT-NUMBER: 20030104584

PGPUB-FILING-TYPE: new

DOCUMENT-IDENTIFIER: US 20030104584 A1

TITLE: Yeast with high biotin-productivity and the preparation method thereof

PUBLICATION-DATE: June 5, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY	RULE-47
<u>Shiuan</u> , David	Kaohsiung		TW	

US-CL-CURRENT: 435/119; 435/254.2, 435/320.1, 435/483

Full	Title	Citation	Front	Review	Classification	Date	Reference	Sequences	Attachments	Claims	QWAC	Draw Desc	Ima
------	-------	----------	-------	--------	----------------	------	-----------	-----------	-------------	--------	------	-----------	-----

[Clear](#)[Generate Collection](#)[Print](#)[Fwd Refs](#)[Bkwd Refs](#)[Generate OACS](#)

Terms	Documents
L7 and L1	1

Display Format:

[Change Format](#)[Previous Page](#)[Next Page](#)[Go to Doc#](#)

Hit List

[Clear](#)[Generate Collection](#)[Print](#)[Fwd Refs](#)[Bkwd Refs](#)[Generate OACS](#)

Search Results - Record(s) 1 through 3 of 3 returned.

☐ 1. Document ID: US 20030104584 A1

Using default format because multiple data bases are involved.

L3: Entry 1 of 3

File: PGPB

Jun 5, 2003

PGPUB-DOCUMENT-NUMBER: 20030104584

PGPUB-FILING-TYPE: new

DOCUMENT-IDENTIFIER: US 20030104584 A1

TITLE: Yeast with high biotin-productivity and the preparation method thereof

PUBLICATION-DATE: June 5, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY	RULE-47
<u>Shiuan</u> , David	Kaohsiung		TW	

US-CL-CURRENT: 435/119; 435/254.2, 435/320.1, 435/483

Full	Title	Citation	Front	Review	Classification	Date	Reference	Sequences	Attachments	Claims	KWIC	Draw Desc	Ima
------	-------	----------	-------	--------	----------------	------	-----------	-----------	-------------	--------	------	-----------	-----

☐ 2. Document ID: US 20030003549 A1

L3: Entry 2 of 3

File: PGPB

Jan 2, 2003

PGPUB-DOCUMENT-NUMBER: 20030003549

PGPUB-FILING-TYPE: new

DOCUMENT-IDENTIFIER: US 20030003549 A1

TITLE: DETECTION OF HUMAN HEPATITIS B VIRUS SURFACE ANTIGEN MUTANTS BY SPECIFIC AMPLIFICATION AND ITS APPLICATION ON GENE CHIP

PUBLICATION-DATE: January 2, 2003

INVENTOR-INFORMATION:

NAME	CITY	STATE	COUNTRY	RULE-47
OON, CHONG-JIN	SINGAPORE		SG	
CHEN, WEI-NING	SINGAPORE		SG	
LEONG, AI-LIN	SINGAPORE		SG	
KOH, <u>SHIUAN</u>	SINGAPORE		SG	

US-CL-CURRENT: 435/110

Full	Title	Citation	Front	Review	Classification	Date	Reference	Sequences	Attachments	Claims	KWIC	Draw Desc	Ima
------	-------	----------	-------	--------	----------------	------	-----------	-----------	-------------	--------	------	-----------	-----

☐ 3. Document ID: US 6759193 B2

L3: Entry 3 of 3

File: USPT

Jul 6, 2004

TITLE: Detection of human hepatitis B virus surface antigen mutants by specific amplification and its application on gene chip

DATE-ISSUED: July 6, 2004

INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Oon; Chong-Jin	Singapore			SG
Chen; Wei-Ning	Singapore			SG
Leong; Ai-Lin	Singapore			SG
Koh; Shiuan	Singapore			SG

US-CL-CURRENT: 435/6; 435/5, 435/91.2, 536/23.72, 536/24.3, 536/24.33

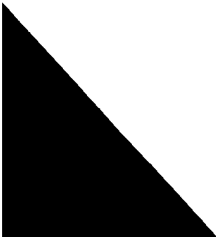
Full	Title	Citation	Front	Review	Classification	Date	Reference			Claims	KWIC	Draw Desc	Ima
------	-------	----------	-------	--------	----------------	------	-----------	--	--	--------	------	-----------	-----

Clear	Generate Collection	Print	Fwd Refs	Bkwd Refs	Generate OACS
-------	---------------------	-------	----------	-----------	---------------

Terms	Documents
L2 and L1	3

Display Format:

[Previous Page](#) [Next Page](#) [Go to Doc#](#)



GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2005, 13:24:44 ; Search time 5408 Seconds
(without alignments)
10644.378 Million cell updates/sec

Title: US-09-752-957F-1
Perfect score: 1188
Sequence: 1 atgtcgtttattatgactgc.....aggttgcatcggtgttga 1188

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1188	100.0	5278	8 AF212161	Candida u
2	519.4	43.7	4068	8 SCTGR286C	733071 S.cerevisia
3	519.4	43.7	35400	8 SCCHIIRA	X94332 S.cerevisia
4	475.8	40.1	110000	8 CR382121_05	Continuation (6 of
5	473.6	39.9	1847	8 SCBIO2GNA	X72701 S.cerevisia
6	458.8	38.6	110000	8 CR382130_18	Continuation (19 of
7	451.6	38.0	110000	8 CR382134_04	Continuation (5 of
8	414.4	34.9	1263	6 AR549314	AR549314 Sequence
9	343.2	28.9	4005	8 SPC320	AL022245 S.pombe c
10	343.2	28.9	40880	8 SPCCL235	AL031764 S.pombe c
11	332.2	28.0	847	11 CMS06K9X	AL402619 T3 end of
12	328.2	27.6	3083	6 AR075303	AR075303 Sequence
13	328.2	27.6	3083	6 I66541	I66541 Sequence 1
14	328.2	27.6	3083	6 AR306523	AR306523 Sequence
15	328.2	27.6	3083	8 SCZU01	X63612 S.cerevisia
16	327.2	27.5	1091	8 SPA224930	AJ224930 Schizosac
17	318.2	26.8	1447	8 AK101236	AK101236 Oryza sat
18	307.2	25.9	1440	8 AK059296	AK059296 Oryza sat
19	304	25.6	1137	6 CQ805820	CQ805820 Sequence

20	304	25.6	1137	6 AX505884	AX505884 Sequence
21	304	25.6	1296	8 ATTHSEACA	I34413 Arabidopsis
22	304	25.6	1332	8 ATU31806	U31806 Arabidopsis
23	304	25.6	1351	6 AR029503	AR029503 Sequence
24	304	25.6	1351	6 AR034920	AR034920 Sequence
25	302.4	25.5	1307	8 BT003848	BT003848 Arabidops
26	300.8	25.3	1342	8 AY088152	AY088152 Arabidops
27	292.6	24.6	1041	6 A11523	A11523 E.coli Bio
28	255.4	21.5	343550	1 AP003587	AP003587 Nostoc sp
29	243.4	20.5	1032	8 BT009521	BT009521 Triticum
30	236	19.9	1017	6 CR320148	CR320148 Sequence
31	231.2	19.5	348672	1 CR378670	CR378670 Photobact
32	229.4	19.3	110000	1 AE017354_16	Continuation (17 o
33	229.4	19.3	110000	2 AP006494_5	Continuation (6 of
34	228.8	19.3	1421	1 AF239257	AF239257 Acinetoba
35	228.8	19.3	349174	1 AB063522	AB063522 Wiggleswo
36	226.4	19.1	210710	6 AX770902	AX770902 Sequence
37	226.4	19.1	349071	1 BX571864	BX571864 Phototrab
38	224.6	18.9	110000	1 CR628337_17	Continuation (18 o
39	223.6	18.8	5729	6 CQ872877	CQ872877 Sequence
40	223.6	18.8	11107	1 U32783	U32783 Haemophilus
41	223.6	18.8	349980	6 CQ873072	CQ873072 Sequence
42	223	18.8	110000	1 CR628336_15	Continuation (16 o
43	215.2	18.1	326050	1 BX321864	BX321864 Nitrosomo
44	215	18.1	348178	1 BX572092	BX572092 Prochloro
45	213.8	18.0	4000	1 AF250768	AF250768 Unculture

ALIGNMENTS

RESULT 1
AF212161

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AF212161
Candida utilis biotin synthase (Bio2) gene, complete cds.
AF212161.1 GI:11596106
Pichia jadinii
Pichia jadinii
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 5278)
Cloning and sequencing of Candida utilis biotin synthase
Unpublished
Hong, Y.R. and Shuan, D.
2 (bases 1 to 5278)
Direct Submission
Submitted (05-DEC-1999) Graduate Institute of Biochemistry, No.
100, Shih-Chuan 1st Road, Kaohsiung, Taiwan

FEATURES
source

1..5278
/organism="Pichia jadinii"
/mol_type="genomic DNA"
/db_xref="taxon:4903"
<3915..>5102
/gene="Bio2"
<3915..>5102
/gene="Bio2"
/product="biotin synthase"
3915..5102
/gene="Bio2"
/codon_start=1
/product="biotin synthase long form"
/protein_id="AAG38483.1"
/db_xref="GI:11596107"

gene

mRNA

CDS

TMKBAEQVMCFMAGCNAIFTGKMLTTMCGWDEDKAMLAUKWGLKPMBSFKYKPREVA
FGA"
4044..5102
/gene="Bio2"
/codon_start=1
/product="biotin synthase short form"
/protein_id="AAG38484.1"
/db_xref="GI:11596108"
/translation="MEEVASQEKPNPLEYALSVKTPVNTWKBEIKAIYDPLMDLM
HYAQVHRFRFOQPSVEQVLTLMNITGGCTEDCKYCAOSQRYNTGVKAEIRIIQVDEVI
EAAKEAKANGSTRFCWGAWEKMRKSNLKIEMITAVHDLGMSCLVTLGMDKQD
ATELKSAGLTYNHNIDTYKHPKVI STRSPFDRLKTFKNVQSGSLCACTGGILG
ETQEDRVFLYLTAMDQHPESLPINRLVPIKGTQPMYEEVKNQVDEIVRTIATAR
LVMPISIIRLAAGRVTMEAEQVMCFMAGCNAI FTGKMLTTMCGWDEDKAMLAUKW
LKPMBSFKYKPREVAFGA"

ORIGIN

Query Match 100.0%; Score 1188; DB 8; Length 5278;
Best Local Similarity 100.0%; Pred. No. 1.5e-308;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTCGTTTATATGACTGCTATTAGTTCGCGATTGCTTTCACCTCTAGAGTAGCT 60
Db 3915 ATGTCGTTTATATGACTGCTATTAGTTCGCGATTGCTTTCACCTCTAGAGTAGCT 3974

Qy 61 TCTAGGCTACTCTGCGACAGGCTCTACTGCTGCTGCGAGATCTTGGAGATGTTTC 120
Db 3975 TCTAGGCTACTCTGCGACAGGCTCTACTGCTGCTGCGAGATCTTGGAGATGTTTC 4034

Qy 121 ACGGAACAAATGGAAGAGTGGCTTTCACAGGAGAAAGACCAACCCATTGGAATATGCA 180
Db 4035 ACGGAACAAATGGAAGAGTGGCTTTCACAGGAGAAAGACCAACCCATTGGAATATGCA 4094

Qy 181 TTGTCAGTGAAGACACCACTGACCTGACCAAGAAAGAAATTAAGCTTATATGAC 240
Db 4095 TTGTCAGTGAAGACACCACTGACCTGACCAAGAAAGAAATTAAGCTTATATGAC 4154

Qy 241 ACACCACTCTGAGCTGATGACCTGATGCTCAGGTGCAACAGAGGTTCCAAACACT 300
Db 4155 ACACCACTCTGAGCTGATGACCTGATGCTCAGGTGCAACAGAGGTTCCAAACACT 4214

Qy 301 TCAGAGGTTCAATTTGCACTCTTATGAATATCAAACTGGTGGTTGACCGAGGACTGT 360
Db 4215 TCAGAGGTTCAATTTGCACTCTTATGAATATCAAACTGGTGGTTGACCGAGGACTGT 4274

Qy 361 AAGTACTGTGCCAATCAGAGGTTTACACACTGCTGTCAGGCTGAAAGAAATCATCCA 420
Db 4275 AAGTACTGTGCCAATCAGAGGTTTACACACTGCTGTCAGGCTGAAAGAAATCATCCA 4334

Qy 421 GTTGATGAGGTGATTTGAAGCTGCAAGGAGGCAAGGCAATGGATCTACAAGGTTCTGT 480
Db 4335 GTTGATGAGGTGATTTGAAGCTGCAAGGAGGCAAGGCAATGGATCTACAAGGTTCTGT 4394

Qy 481 ATGGGTGCTGCTGGAGAGAGATGAAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAG 540
Db 4395 ATGGGTGCTGCTGGAGAGAGATGAAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAG 4454

Qy 541 ATGATCACTGCTGTCTGATGACCTTGGAAATGGAGAGTTGTGTCAACCTGGGAATGGTTGAT 600
Db 4455 ATGATCACTGCTGTCTGATGACCTTGGAAATGGAGAGTTGTGTCAACCTGGGAATGGTTGAT 4514

Qy 601 AAAGACCAAGCCACTGAATTTGAAGTCTGCTGGTTGACGGCTGACCAACCAATCAATGAT 660
Db 4515 AAAGACCAAGCCACTGAATTTGAAGTCTGCTGGTTGACGGCTGACCAACCAATCAATGAT 4574

Qy 661 ACTTACAAGGAACACTATATCAAGGTGATCTCCCAAGAGGCTTTGATGATAGATTGAAA 720
Db 4575 ACTTACAAGGAACACTATATCAAGGTGATCTCCCAAGAGGCTTTGATGATAGATTGAAA 4634

Qy 721 ACATTAACAAAGCTTCAAGGATCTGGAATTAAGGCAATGACAGGTGGTATCTTGGTCTT 780
Db 4635 ACATTAACAAAGCTTCAAGGATCTGGAATTAAGGCAATGACAGGTGGTATCTTGGTCTT 4694

Qy 781 GGTGAGACCCAGAGGACCGTGTATCTTTCTCTACCTTTGGCCACAAATGGATCAGCAT 840
Db 4695 GGTGAGACCCAGAGGACCGTGTATCTTTCTCTACCTTTGGCCACAAATGGATCAGCAT 4754

Qy 841 CCAGAGTCTCTTCCAAATCAACAGACTGGTCCCAATCAGGGCAGCCCAATGTATGAAGAA 900
Db 4755 CCAGAGTCTCTTCCAAATCAACAGACTGGTCCCAATCAGGGCAGCCCAATGTATGAAGAA 4814

Qy 901 GTTAAAGAACAGCAAGTTGAAGTTGATGAGATTGTGAGAACCATTTGCTACTGCAAGATTG 960
Db 4815 GTTAAAGAACAGCAAGTTGAAGTTGATGAGATTGTGAGAACCATTTGCTACTGCAAGATTG 4874

Qy 961 GTCATGCCAACGCTCTATTATCAGATTGGCTCAGGAAGATATACAATGAAGAGGACAGAA 1020
Db 4875 GTCATGCCAACGCTCTATTATCAGATTGGCTCAGGAAGATATACAATGAAGAGGACAGAA 4934

Qy 1021 CAGTGATGCTCTTCATGGCTGGTGTAAATGCCATCTTCACAGGTAAAGAAATGCTCACA 1080
Db 4935 CAGTGATGCTCTTCATGGCTGGTGTAAATGCCATCTTCACAGGTAAAGAAATGCTCACA 4994

Qy 1081 ACAATGTGTAACGCTGGGATGAGGATAAAGCCCATTTGGCTAAATGGGGTCTGAAACCA 1140
Db 4995 ACAATGTGTAACGCTGGGATGAGGATAAAGCCCATTTGGCTAAATGGGGTCTGAAACCA 5054

Qy 1141 ATGAGAGATTTCAAATACAAACCAAGGAGGTTGCATTCGGTGTGTA 1188
Db 5055 ATGAGAGATTTCAAATACAAACCAAGGAGGTTGCATTCGGTGTGTA 5102

RESULT 2
SCYGR286C/c 4068 bp DNA linear PLN 11-AUG-1997
LOCUS S.cerevisiae chromosome VII reading frame ORF YGR286C.
DEFINITION
ACCESSION Z73071 Y13135
VERSION Z73071.1 GI:1323522
KEYWORDS
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE 1 (bases 1 to 4068)
AUTHORS Voet, M. and Volckaert, G.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4068)
AUTHORS MIPS.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-1996) Data collected by MIPS on behalf of the
European yeast chromosome VII sequencing project. MIPS at the
Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
Martinsried, FRG; E-mail: Mewes@mips.emblnet.org
FEATURES
source Location/Qualifiers
1..4068
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/chromosome="VII"
complement (660..1787)
/gene="BIO2"
complement (660..1787)
/gene="BIO2"
/note="ORF YGR286C"
/codon_start=1
/protein_id="CAA97318.1"
/db_xref="GI:1323523"
/db_xref="GOA:P32451"
/db_xref="SGD:S0003518"
/translation="MMSTIYRHLSTARPAATKYATNAAVKSTTASSEASTIGALQYAL
SLDPSHSWTKSLKELIYHTPLLELTHAAQLQHRKWHDPKVOICLTLMNISKGSSED
CKYCAQSRNDTGLKAEKWKVDEIKEAEKRGSTFRFCGLGAARMDKRGKAMKR
IQEMVTKVNDMGLETCTVLTGMDVDOAKQKADAGLTAVNHNIDTSRHSYKVIITRTY
DDRIQTIKNVOESGIKACTGILGISEDDHIGFIYTLNMSPHPSLPINRLVAIK
GTPMAEELADPKSKKLOPDEILRTIATARIIVMPKAILINLAAGRYTNKMETQFVCPMAG

CNSIFTGKKMLTTMCNGWDEPKAMLAKWGLOPMEAFKYDRS"

RESULT 3	SCCHVIIRA	35400 bp	DNA	linear	PLN 14-APR-1997
LOCUS	SCCHVIIRA	35400 bp	DNA	linear	PLN 14-APR-1997
DEFINITION	S.cerevisiae DNA for fragment from chromosome VII.				
ACCESSION	X94332				
VERSION	X94332.1	GI:1262143			
KEYWORDS					
ORGANISM	Saccharomyces cerevisiae (baker's yeast)				
SOURCE	Saccharomyces cerevisiae				
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.				
AUTHORS	Volckaert, G., Voet, M. and Robben, J.				
TITLE	Sequence analysis of a near-subtelomeric 35.4 kb DNA segment on the right arm of chromosome VII from Saccharomyces cerevisiae carrying the MAL1 locus reveals 15 complete open reading frames, including ZUO1, BGL2 and B5I2 genes and an ABC transporter gene				
JOURNAL	Yeast 13 (3), 251-259 (1997)				
MEDLINE	97245295				
REFERENCE	9090054				
AUTHORS	Volckaert, G.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-DEC-1995) G. Volckaert, Laboratory of Gene Technology, Catholic University of Leuven, Willem de Croylaan 42, B-3001 Leuven, BELGIUM				
FEATURES	Location/Qualifiers				
source	1..35400				
	/organism="Saccharomyces cerevisiae"				
	/mol_type="genomic DNA"				
	/strain="FY1679"				
	/db_xref="taxon:4932"				
	/chromosome="VII"				
	/map="right arm"				
	/clone="cosmid pEGH452"				
	/clone_lib="H.tetellin (Louvain-la-Neuve, Belgium)"				
ORIGIN					
Query Match	43.7%;	Score 519.4;	DB 8;	Length 35400;	
Best Local Similarity	70.9%;	Pred. No. 1.4e-128;			
Matches	707;	Conservative	0;	Mismatches 281;	Indels 9; Gaps 1;
Qy	173	AATATGCATTGTTCAGTGAAGACACCAAGTCAACCTCGGACCAAGAGAAATTTAAAGCTA	232		
Db	15763	AATATGCATTGTCTTTAGATGAACCAAGTCATTCTGTGGACAAAATCGCAATTTAAAGAAA	15822		
Qy	233	TATATGACACCACTCTGAGCTTGATGTCATATGCTCAGGTGCACACAGAGGTTCC	292		
Db	15823	TTTATCATACCCCACTGCTCGAACTTACTCATGTCAGACCAATTTGCAGCACAGAAAAGTGGC	15882		
Qy	293	AACAACCTTCAGAGGTTCAATTGTGCACCTCTTATGAATATCAAAACTTGTTGTTCTACCG	352		
Db	15883	ACGATCCAACCAAGTGCATTTGTGCACATGATGAACATCAATCTGGTGGTTCGTG	15942		
Qy	353	AGGACTGTAAGTACTGTGCCCCAATCACAGCGTTACAACACTGGTCTCAAGGCTGAAGAA	412		
Db	15943	AGGACTGTAAGTATTGTGGCGAGTCTTCGAGAAACGATACCGGTCTAAGGCTGAGAAA	16002		
Qy	413	TCATCCAAGTTCATGAGGTGATTGAAGCTGCAGAGGCGCAAGGCCAATGATCTCAAA	472		
Db	16003	TGGTTAAGTGGATGAAGTGATTAAAGAGCGAAGAGGCCAAAAGAACCGGATCTCTA	16062		
Qy	473	GGTTCTGTATGGGTGCTGCTTGGAGAGAGATGAAAGGTAGAAAGTCAAACTTTGAAGAAA	532		
Db	16063	GATTTCTGCCTAGTCTGTCATGAGAGACATGAAGAGGTCGTAATACGCCATGAAGAA	16122		
Qy	533	TCAAAGAGATGATCACTGCTTCATGACCTTGGAAATCGAGAGTTGTGTCAACCTCGGAA	592		
Db	16123	TTTCAGGAAATGTTGACCAAGTGAATGATATGTTGGGCTAGAAACCGTGTGTTACTTTAGTA	16182		

ORIGIN	Query Match Best Local Similarity Matches	43.7%; 70.9%; 707;	Score 519.4; Pred. No. 1.5e-128; Conservative 0;	DB 8; Mismatches 281; Indels 9;	Length 4068; Gaps 1;
Qy	173	AATATGCATTGTCAGTGAAGACACCAAGTCAACACTGGACCAAGAGAAATTTAAAGCTA	232		
Db	1666	AATATGCATTGTCATTAGTGAACCAAGTCAATTGTTGGACAAAATCGCAATTTAAAGAAA	1607		
Qy	233	TATATGACACACCACTCATGAGCTTGTAGTCACTATGCTCAGGTGCAACACACAGAAGGTTC	292		
Db	1606	TTTATCATACCCCACTGCTCGAATTACTCATGACGCAATTTGCACACAGAAGGTGCG	1547		
Qy	293	AACAACCTTCAGAGGTTCCAAATTGTGCACCTCTTATGAATATCAAACTGGTGGTTGTACCG	352		
Db	1546	ACGATCCAAACCAAGTGCAAATTGTGCACATTGATGAACATCAAAATCTGGTGGTTGTTCTG	1487		
Qy	353	AGGACTGTAACTATGTGCCCAATCAACAGCTTACACACTGGTGTCAAGGCTCAAGAA	412		
Db	1486	AGGACTGTAACTATGTGCGCAGTCTTCGAGAAAACGATACCGGTCTTAAAGGCTCAGAAA	1427		
Qy	413	TCATCCAAGTTGATGAGGTGATTCAAGCTGCAAAAGGAGGCAAGCGCAATGGAATCTCA	472		
Db	1426	TGGTTAAAGTGGATGAAGTGAATTAAGGCGCAAGAGGCGCAAAAGAAACGGATCTACTA	1367		
Qy	473	GGTTCTGTATGGGTGCTGCTTTGGAGAGAGATGAAGGTGAGAAAGTCAAACTTTGAAGAAA	532		
Db	1366	GATTTCTGCTAGGTGCTGCATGAGAGAGACATGAAGGTTCGTAAATCAGCCATGAAAA	1307		
Qy	533	TCAAGAGATGATCAGTCTGCTTCATAGCTTGGAAATGGAGAGTTGTACACCTGGAA	592		
Db	1306	TTCAGGAAATGGTGAACCAAGTGAATGATATGGGGCTAGAAACCGTGTGTTACTTTTAGGTA	1247		
Qy	593	TGGTTGATAAAGACCAAGCCACTCAATTTGAAAAGTGTCTGGGTTGACGGCGTCAACACATA	652		
Db	1246	TGGTTGATCAAGATCAAGCAAGCAATTTGAAAGATGCAAGTTTGACTGCATACACACATA	1187		
Qy	653	ACATTGATCTTCAAGGAACACTATCCAAAGGTGATCTCCCAAGAAAGCTTTGATGATA	712		
Db	1186	ACATCGACACTTCAGAGAACACTATATAGTAGGTTCATCAACACAGAAACCTCAAGACGACA	1127		
Qy	713	GATTTGAAAACATTCANAAACGTTCAAGGATCTGGATTAAAGGCAATGACAGGTGGTATTC	772		
Db	1126	GGTTACAGACCATCAAGAATGTCCAGAACTCGGAATTAAGAGCTGTACCGGTGGTATTT	1067		
Qy	773	TTGTCCTTGGTGAGACCCAAAGAGACCGTGTATCTTTCCCTCTACACCTTGGCCACAAATGG	832		
Db	1066	TGGTCTCGTGAAAGCGCAAGACCATATAGGATTCATCACATATCCAAATATGT	1007		
Qy	833	ATCAGCATCAGAGTCTCTTCCAAATCAACAGACTGGTCCCAATCAAGGGCAACGCCAATGT	892		
Db	1006	CTCCTCATCTGAGTCCCTACCAATTAATAGACTAGTTGCTATCAAAAGGACCTCCAAATGG	947		
Qy	893	ATGAAGAAGTT-----AAGAAACAGCAAGTTGAAGTTGATGAGATGTCAGAAACCA	943		
Db	946	CTGAGGAACCTTGGCGATCCAAAGAGTAAAGTTGCAATTCGACGAATTTTGAGAAACCA	887		
Qy	944	TTGCTACTGCAAGATTGGTCAATGCCACGCTATTATTCAGATTGGCTGCAAGGAGATATA	1003		
Db	886	TTGCCACAGCGAATAAGTTATGCCAAGGCCAATTAAGACTTGGCGCTGGTCGTTATA	827		
Qy	1004	CAATGAAGAGGAGACAGGTGATGCTTCATGGCTGGTGTGTAATGCCATCTTCACAG	1063		
Db	826	CAATGAAGAAACAGAGCAATTTGTCTGTTTCATGGCAGGTGTATAAGTATCTTCACCG	767		
Qy	1064	GTAAAGAAATGCTCACAAACAATGTAAACGGCTGGGATGAGGATAAAGCCATCTTCGGCTA	1123		
Db	766	GTAAAGAAATGCTGACGACAAATGTATAACGGTTGGGACGAGACAGGCATCTTCGGCTA	707		
Qy	1124	AATGGGCTGCAAAACCAATCGAGGTTTCAAAATACAA	1160		

593 TGGTTGATTAAGACCAAGCCACTGAATGAAAAGTGTGGGTGTCAGCGGTACCAACATA 652
16183 TGGTTGATCAAGATCAAGCAAGCAATTTGAAGATGCAAGTTTGACTGCATACCAACATA 16242
653 ACATTGATTAACAGGAAACACTATCCAAAGGTGATCTCCACAAAGAGCTTTGATGATA 712
16243 ACATCGACACTTCAGAGAAACACTATAGTAGTATAGTATCACCACGAGAACCTTACGACGACA 16302
713 GATTGAAACATTCAAAACGTTCAAGGATCTGGAATTAAGGCAATGCAACAGGTGATTC 772
16303 GGTTCACAGCAATCAAGAAATGTCAGAAATCTGGAATTAAGGCAATGCAACAGGTGATTC 16362
773 TTGCTCTTGGTGAGACCAAGAGGACCGTGTATCTTTCTCTACACCTTGGCCCAATGG 832
16363 TGGGTCTCGGTGAAGCGAAGACGACCATATAGGATTCATCTACACATTAATCCAAATGT 16422
833 ATCAGCATCCAGAGTCTTCCAAATCAACAGACTGGTCCCAATCAAGGGCAGGCAATGT 892
16423 CTCCTCATCTCTGAGTCCCTACCAATTAATAGACTAGTGTCTATCAAGGGGACTCCAAATGG 16482
893 ATGAGAGATT-----AAGAACAGCAAGTTGAAGTTGATGAGATTGTCAGAACCA 943
16483 CTGAGGAACTTGGCGATCCAAAGAGTAAAAAGTTGCAATTCGACGAAATTTTGAAACCA 16542
944 TTGCTACTGCAAGATTGGTCAATGCAACGCTCTATTTATCAGATTGGCTGCGAGAAAGATATA 1003
16543 TTGCCACGCGAGATAGTTATGCCAAAGGCCATTAATGAAGCTTGGCGCTGGTGTATA 16602
1004 CAATGAAAGAGGCAAGACAGGTGATGCTTCAATGCTGCTGTTGTAATGCCATCTTCACAG 1063
16603 CAATGAAAGAAACAGACAAATTTGCTGTTTCAATGCGCAGGTTGTAACAGTATCTTCACCG 16662
1064 GTAAGAAATCTCACAACATGTGTAACGCTGGATGAGGATAAAGCCATGTTGGCTA 1123
16663 GTAAGAAATCTGACGCAATGTGTAACGCTGGGAGCAAGCAAGCAATGTTGGCTA 16722
1124 AATGGGTCTGAAACCAATGGAGAGTTTCAAAATACAA 1160
16723 AATGGGATTGCAACCTATGAGGCAATTAAGTAGCA 16759

RESULT 4

CR382121_05
WPCOMMENT

Sequence split into 11 fragments LOCUS CR382121 Accession CR382121

Fragment Name	Begin	End
CR382121_00	1	110000
CR382121_01	100001	210000
CR382121_02	200001	310000
CR382121_03	300001	410000
CR382121_04	400001	510000
CR382121_05	500001	610000
CR382121_06	600001	710000
CR382121_07	700001	810000
CR382121_08	800001	910000
CR382121_09	900001	1010000
CR382121_10	1000001	1062590

Continuation (6 of 11) of CR382121 from base 500001 (CR382121 Kluyveromyces lactis strain

Query Match 40.1%; Score 475.8; DB 8; Length 110000;
Best Local Similarity 67.8%; Pred. No. 7.4e-117;
Matches 685; Conservative 0; Mismatches 317; Indels 9; Gaps 1;

175 TATGATTTGTCAGTGAAGACACAGTCAACACCTGGACCAAGAGAAATTAAGCTATA 234
46048 TATGCACTATCTTTGAACGAACACAGCATGTGTGGACCAAGAACATTAAGCGAGATT 46107
235 TATGACACCACTCATGAGCTTGTATGCTATGCTCAGGTGCAACAGAGGTTTCAA 294
46108 TATCACGCGCATGATGAATTTGATGCTCAGGCAATTAACATAGGAATGGCAT 46167
295 CAACCTTCAGAGGTTCAATTTGCACTCTTATGAATATCAAACTGGTGGTTGACCGAG 354

Db 46168 GATCTCTACCAAGGTGCAACTGTGTACATTGATGAACATCAAGGCGGTGAGTTTCAGAA 46227
Qy 355 GACTGTAAAGTACTGTGTCCCAATCAACAGCGTTTACAACTGTGTCAAGGCTGAAAGAAATC 414
Db 46228 GATTTGAAGTATTGTGGCGCAATCTTCAAGTTACGATACAGGTCTTTAAAGCAGAGAAGATG 46287
Qy 415 ATCCAAGTTGATGAGGTGATGAGCTGCAAGAGGCAAGGCCAATGGAATCTTACAGG 474
Db 46288 ATATCTGTCGAAGAAGTGAATCAAGGAGGCAAAATTTGCCAAAGAAAATGGGTCCACGAGA 46347
Qy 475 TTCTGTATGGTCTGCTTCGGAGAGAGATCAAGGTTAGAAAGTCAAACTTTGAAGAAAATC 534
Db 46348 TTTTGTCTTGGTCTGCTTCGGAGAGAAATCGGTGTAGAAAATCTGCAATTTGAAGAAGATTT 46407
Qy 535 AAAGAGATGATCACTGCTGTCTCCATGACCTTTGGAAATGGAGAGTTGTGTCAACCTGGGAATG 594
Db 46408 GGAGAATGATTGGAAAGATTAAAGAGATGGGATGGAATCTTGGGTGACATTTGGGAATG 46467
Qy 595 GTTGATAAAGACCAAGCCACTGAATTTGAAAAGTGTGGGTGTAGGGGTGACGGGTACACCAATAC 654
Db 46468 GTGGATGAAACTCAAGCGAAACAGTTTGAAGGACGCTGGTTTGACTGTCTTAAACCAATAC 46527
Qy 655 ATTGATATCTTACAAGGACACTATCCAAAGGTGATCTCCACAGAAAGCTTTTGATGATAGA 714
Db 46528 ATCGATACCTCGAGGGGAACATTAACCAAGGTGATCTTACCAGATCATATGACGATAGA 46587
Qy 715 TTGAAAACATTTCAAAAACGTTTCAAGGATCTGGATTTAAAGCATGCAAGGCTGGTGTATCTTT 774
Db 46588 TTGCAAAACCAATTAATAATGTGCAAGATGCAAGTATTAAGCTTTGCAAGGTGGATCTTG 46647
Qy 775 GGTCTTGGTGAGACCCCAAGAGGACCGGTATCTTTCTCTACACCTTTGGCCACAAATGGAT 834
Db 46648 GGTCTTGGTGAACCTGAAGATGATCACATTTGGGTTCGTTTCCACACTTTTGGCAAAATATGAT 46707
Qy 835 CAGCATCCAGAGTCTCTTCAATCAACAGACTGCTCCCAATCAAGGCGACGCCCAATGTAT 894
Db 46708 CCACATCCGAATCTTTTACCTATTATATAGATTAGTTTCCATTTAAAGGTACACCAATGAT 46767
Qy 895 GAAGAAGTTAAGAA-----CAAGCAAGTTGAAGTTGATGATGATTTGTGAGAACCAAT 945
Db 46768 GAGGAACTTAAGAAGGTCAGCAGTAAAGAAATTTGGAATTCGATGAGATCTTAAAGAACCGTG 46827
Qy 946 GTTACTGCAAGATTGGTCAATGCGTCCAAACGCTTATTATCAGATTGGCTGAGGAGATATACA 1005
Db 46828 GCATAGCCAGAAATTTGAATGCCACAGCGAATTAATCAGATTAGCTGCGAGTTCGTACACG 46887
Qy 1006 ATGAAAGAGGCAAGACAGGTGATGCTTCAATGCTGCTGTTGTAATGCAATCTTCCACAGGT 1065
Db 46888 ATGAAGGAAACTGAACAAATTTCTTATGCTTTCATGCGAGTTGTAATGCCATCTTTGTCGGT 46947
Qy 1066 AAGAAAATGCTCAACAAATGTGTAAAGGCTGGGATGAGGATTAAGGCAATGTTGGCTAAA 1125
Db 46948 AAAAGATGTTTAAACAACTATGTGTACAGGCTGGGATGAGGCAAGCAATGCTTTTCGAA 47007
Qy 1126 TGGGGTCTGAAACCAATGAGAGATTTCAAATACAAACCAAGGGAGGTTGCA 1176
Db 47008 TGGGGTCTACAAACCAATGGAATCGTTTTCATATATCCGCTAATAAGGCTCCA 47058

RESULT 5

SCBIO2GNA

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

SCBIO2GNA 1847 bp DNA linear PLN 07-APR-1994
S.cerevisiae gene for biotin synthetase.

X72701 GI:288356
X72701.1 S.cerevisiae (baker's yeast)

Bio2 gene; biotin synthetase.
Saccharomyces cerevisiae
Saccharomycetales; Ascomycota; Saccharomycetes;
Eukaryota; Fungi; Ascomycota; Saccharomycetales; Saccharomycetes;
1 (bases 1 to 1847)

Zhang, S., Sanyal, I., Bulboacă, G.H., Rich, A. and Flint, D.H.
The gene for biotin synthase from Saccharomyces cerevisiae.

```

cloning, sequencing, and complementation of Escherichia coli
strains lacking biotin synthase
Arch. Biochem. Biophys. 309 (1), 29-35 (1994)
94161552
8117110
2 (bases 1 to 1847)
Zheng,S.
Direct Submission
Submitted (12-MAR-1993) S. Zheng, 16-739 Dept. of Biology,
Massachusetts Institute of Technology, Cambridge, MA -2139, USA
Location/Qualifiers
1. .1847
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="20B-12"
/db_xref="taxon:4932"
/chromosome="VII"
159. .1226
/gene="B102"
159. .1226
/gene="B102"
/codon_start=1
/product="biotin synthetase"
/protein_id="CAA51253.1"
/db_xref="GI:288357"
/db_xref="GOA:P32451"
/db_xref="UniProt/Swiss-Prot:P32451"
/translation="MPQLNRQLHQKLVPGCSTICIVRLTKSFVDKIAIKRNLVSPIT
ARYSCSHNCSHRKHDPTRVQLCTLNNIKSGSCSDCKYCAQSSNRDITGLKAEKWYK
VDEVIKRGRGCKRNGSTRFCLGAARDMKRGSAMKRIQEMTYKNDMDLETCTVLIG
MVPDQAKQLDKADTHAYNNIDTSREHYSKVITRTYDDRLQITKNVQSGIKACTDE
GILGLGSSDDHIGFIYTLNMSHPHESLPINRLVAIKGTPMAEELADPKSKKLQDFE
ILRTIATVATVMPKAIIRLAAGRTYMKETETQEFVCFMAGCNSIFTGKQWLTTIYNGWDE
DKAMIAKAGIQQEFAEAYDS"

```

RESULT 6
CR382130_18
WPCOMMENT

Query Match	Best Local Similarity	Score	DB	Length
CR382130_32	3200001	3310000	DB 8	110000
CR382130_33	3300001	3410000		
CR382130_34	3400001	3510000		
CR382130_35	3500001	3610000		
CR382130_36	3600001	3633272		
Continuation (19 of 37)				
Query Match	38.6%	Score 458.8	DB 8	Length 110000
Best Local Similarity	68.3%	Pred. No. 2.8e-112		
Matches	653	Conservative	0	Mismatches 297
				Indels 6
				Gaps 1
Qy	202	AACACCTGGACCAAGAAATTAAGCTATATATGACACACACCTCATGGACTTCATG	261	
Db	84199	ACCAACTGGACGAGATGAGGTGTCCAAATCTACCAACATGCTTTGATGGATCTTGT	84258	
Qy	262	CACATCTCAGGTGCAACACAGAGGTTCACCAACCTTCAGAGGTTCAATTCGTGCACT	321	
Db	84259	TTCAGGCGCGCCACCGTGCAACACACACCAACCTTCGAGGTTGCAACTTTGCACT	84318	
Qy	322	CTTATGAATATCAAAATCTGGTGTGTACCGAGGACTGAAGTACTGTGCCAATACAG	381	
Db	84319	CTTTACAACATCAAGATGGAGGATGCATGAGGACTGCAAGTACTGGCTCAGTCATCC	84378	
Qy	382	CGTTACAACACTGTGTCAAGGCTGAAGAAATCATCCAAAGTTGATGAGGTGATTGAAGCT	441	
Db	84379	AGATACACTACTGTGTGTGAGCGCGAAGAGGACGTCACAGCCAGCTGAGGTATTGGCGCT	84438	
Qy	442	GCAAGGAGGCAAGGCCNATGGATCTACAGGTTCTGTATGGGTGCTGTTGGAGAGAG	501	
Db	84439	GCAAGGAGGCAAGGCCNATGGTTCTACGAGATCTGTATGGGAGCGCTGGCGAGAC	84498	
Qy	502	ATGAAAGGTAGAAAGTCAAACTTCAAGAAATCAAGAGATGATCATCTGTGTCCATGAC	561	
Db	84499	ATGAAAGGTGCAAGAGACGCCCTTAAGNACGTCAAGGAGATGATTGGTGTGACAGAG	84558	
Qy	562	CTTGGAAATGAGAGATTTGTTCACCTTGGGAATGTTGATTAAGACCAAGCCACTGAATTG	621	
Db	84559	CTCGCATGGAATCTCTGCTGACTCTCGGAATGTTGACAGGCTCAGGCTGAAGAGCTA	84618	
Qy	622	AAAAGTCTGGTTGACGGGTCAACCAATCAACATTCATCTTCAAGGAAACACTATCCA	681	
Db	84619	CGAAGACTGGACTCCTGCTGCTCAACCAATCAACATTCATCTTCAAGGAAACACTACCT	84678	
Qy	682	AAGGTGATCTCCACAAGAAGCTTTGATGATAGATTGAAAACATTCAAAAAGTTCAAGGA	741	
Db	84679	AAGGTGATCACCCTCGAATCTTCGATGAGCGCTCTCAGACTATTGAAGCTTAGAGAA	84738	
Qy	742	TCTGGATTAAGGATGACAGGTGGTATTTCTGGTCTTTGGTGAGACCCAGAGGACCGT	801	
Db	84739	GCAGGCGTCAGCGTGTCTCTGTTGGTATTCTTGGTCTGGGTGAAACAGCTGATGACCAC	84798	
Qy	802	GTATCTTCTCTACACTTGGGCGACATGATCAGGATCAGGATCTCTTCCAATCAAC	861	
Db	84799	GTTGGATTCCTTGGGCTCTGGCTACTATGCTCTCTATCTGAGTCCCTTCTCTATCAAC	84858	
Qy	862	AGACTGTGCCAATCAAGGGGACGCCAATGATGAAGA-----AGTTAAGAAACAAGCAA	915	
Db	84859	CGTCTTGTCCCATCAAGGGGACACTCTTGGAGAGACCTCGCCATGAGGAGAGAGCAA	84918	
Qy	916	GTTGAAATTTGATGAGATTTGTGAGAACCAATTGCTACTGCAAGATTTGGTCAATGCTCT	975	
Db	84919	CTCACCTATGACACTGTTGTTGCAAGTGTCTGCTACTGCTGAGTGGTGTGCTGGCTCT	84978	
Qy	976	ATTATCAGATTTGGCTGAGGAGATATACATGAAGGCGAGAACAGGTGATGTGCTTC	1035	
Db	84979	AAAATCAGACTTGTCTGTGAGACGATACCAATGAAGGAACTGAGCAGGCGCATGTGCTTC	85038	
Qy	1036	ATGCTGTTGTTAATGTCATCTTCAAGGTAAGAAAATGCTCAACAATGTTGTAACGGC	1095	
Db	85039	ATGCTGAGGCAAGCCACTTTACAGGCAAGAGATGCTGACTACCATGTGCAACGGC	85098	
Qy	1096	TGGATGAGGATTAAGCCATGTTGGCTAAATTTGGGCTGTGAACCAATGAGAGTTT	1151	

Db 85099 TGGGACAGGANTAGGCCATGCTGGAGNATGTGGGCACTGAAGCCCATGAAGAGCTT 85154

RESULT 7
CR382134_04/c
WPCOMMENT

Sequence split into 14 fragments LOCUS CR382134 Accession CR382134

Fragment Name	Begin	End
CR382134_00	1	110000
CR382134_01	100001	210000
CR382134_02	200001	310000
CR382134_03	300001	410000
CR382134_04	400001	510000
CR382134_05	500001	610000
CR382134_06	600001	710000
CR382134_07	700001	810000
CR382134_08	800001	910000
CR382134_09	900001	1010000
CR382134_10	1000001	1110000
CR382134_11	1100001	1210000
CR382134_12	1200001	1310000
CR382134_13	1300001	1349926

Continuation (5 of 14) of CR382134 from base 400001 (CR382134 Debaryomyces hansenii chr

Query Match	38.0%; Score 451.6; DB 8; Length 110000;
Best Local Similarity	65.9%; Pred. No. 2.4e-110;
Matches 696; Conservative 0; Mismatches 339; Indels 21; Gaps 2;	
QY 138 AGTGGCTTCACAGGAGAAGGCCAACACCCATTCGGAATATGCATTGTGCTAGTGAAGACACC 197	
Db 31173 AGACACTGGCCACAGCTCAATATGAAGGCTATTGAAATTAGCATTTACAGACCCCTAGACC 31114	
QY 198 AGTCAACACCTGGACCAAGAAGAAATTAAGACTATATATGACACACACCACCTCATGGACTT 257	
Db 31113 GTTGAACTCATGACCAAGACCAATCCGTCGAGATATACAATACGCCGCTAATGGAACCT 31054	
QY 258 GATSCACTATGCTCAGGTGCAACACAGAGGTTCCAAACACCTTCAGAGGTTCAATTGTG 317	
Db 31053 AGTTTTTCAATCTCAAAATTCAAACATAGGAAATATCAACACCTCAAGAGGTTCAAGTTATG 30994	
QY 318 CACTCTTTATGAATATCAAACTGGTGTGTACCGAGACTGTAACTACTGTGCCCAATC 377	
Db 30993 TACTTTGTTATCAATCAAGACTGGCGTTGTACCGAGATTGCAAGATTATGTGCGCAATC 30934	
QY 378 ACAGCGTTACAACACTGGTGTCAAGGCTGAAAGAAATCATCAAGTTGATGAGGTGATTGA 437	
Db 30933 GACCAGATATGATACTGGAACCAAGGCTGAGAGGATGATTTCAGTAGATGAAGTTATGAA 30874	
QY 438 AGTGTCAAAGGAGGCAAGGCCAATGATCTACAAGGTTCTGTATGGTGTGCTGTGAG 497	
Db 30873 GGCTGCTAGAACTGCAAAAGCAAAATGGATCTACTAGATTCTGTATGGGAGCTGATGGAG 30814	
QY 498 AGAGATCAAAAGGTAGAAAGTCMAACTTGAAGAAAT---CAAAGAGATGATCACTGCTGT 554	
Db 30813 AGACATGACAGGTGCTAAGTCAGATTGAAGCGTATTTCCTCAAAATGGTTACCCAAATTA 30754	
QY 555 CCATGACCTTTGGAATGGAGAGTTGTGTCAACCCTGGGAATGGTTGATTAAGACCAAGCCAC 614	
Db 30753 TCAGAGCTAGGATGGAACATGTGTGACACTAGGAATGGTTGATAAGAAACAGGCGGC 30694	
QY 615 TGAATTGAAAGTCTGGTTGACGGCGTACAACATCAATGATCTTACAGAGGAACA 674	
Db 30693 AGACTTAGCTGATCAGGTTTGACTGGCTACAACGATAACATCATCATCTAGAGAGCA 30634	
QY 675 CTATCCAAAGGTGATCTCCAAGAAGCTTTGATGATAGATTGAAMACATTCAAAAACGT 734	
Db 30633 TTATCCAAAGGTATTTCAACTAGAGTTACGATGAGAGGTTGCAAACATCAGGAATGT 30574	
QY 735 TCAAGGATCTGGATTAAAGGCATGCACAGGTGGTATTCTTGGCTTGGTGAGACCAAGA 794	
Db 30573 TCAAGATGCTGGTATTAAAGCATGCTCTGGAGGTATCCTTGGACTTGGTGAACCTGAGGA 30514	
QY 795 GGACCGGTGATCTTTTCCCTCTACACCTTGGGCCAATGGATCAGCATCCAGATCTCTTCC 854	

Squares, R., Squares, S., Stevens, K., Taylor, K., Taylor, R.G., Tivey, A., Walsh, S., Warren, T., Whitehead, S., Woodward, J., Volckaert, G., Aert, R., Robben, J., Grymonprez, B., Weijens, I., Vanschaels, E., Rieger, M., Schafer, M., Muller-Auer, S., Gabel, C., Fuchs, M., Dusterhoft, A., Fritz, C., Holzer, E., Moesti, D., Hilbert, H., Borzym, K., Langer, I., Beck, A., Lehrach, H., Reinhardt, R., Pohl, T.M., Eger, P., Zimmermann, W., Wedler, H., Wambutt, R., Purnelle, B., Goffeau, A., Cadieu, E., Dreano, S., Gloux, S., Lelaure, V., Mortier, S., Galibert, F., Aves, S.J., Xiang, Z., Hunt, C., Moore, K., Hurst, S.M., Lucas, M., Rochet, M., Galliard, C., Tallada, V.A., Garzon, A., Thode, G., Daga, R.R., Cruzado, L., Jimenez, J., Sanchez, M., del Rey, F., Benito, J., Dominguez, A., Revuelta, J.L., Moreno, S., Armstrong, J., Forsburg, S.L., Cerutti, L., Lowe, T., McCombie, W.R., Paulsen, I., Potashkin, J., Shpakovski, G.V., Ussery, D., Barrall, B.G. and Nurse, P.

The genome sequence of *Schizosaccharomyces pombe*
Nature 415 (6874), 871-880 (2002)

2 (bases 1 to 40005)

Wood, V., Rajandream, M.A., Barrell, B.G., Wedler, H., Wambutt, R. and Wedler, E.

Direct Submission
Submitted (24-MAR-1998) European Schizosaccharomyces genome sequencing project, Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: pombe@sanger.ac.uk and ACGWA GmbH, Glienicker Weg 185, D-12489 Berlin, Germany

On Apr 14, 1999 this sequence version replaced gi:2995362.

Notes:
Details of *S. pombe* sequencing at the Sanger Institute are available on the World Wide Web.
URL, <http://www.genedb.org/genedb/pombe/index.jsp> (URL, http://www.sanger.ac.uk/Projects/S_pombe/)
CDS are numbered using the following system eg SPAC5H10.0lc.SP (S. pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c (complementary strand). However, clones may have been reorientated since the original submission, therefore the complementary strand notation may be invalid for strand inference. IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once.

Location/Qualifiers
1..40005
/organism="Schizosaccharomyces pombe"
/mol_type="genomic DNA"
/strain="972h"
/db_xref="taxon:4896"
/chromosome="III"
/map="IIIL"
/clones="cosmid c320"
1..7687
/note="nominal overlap with cosmid SPCC330, EM:AL031603 S. pombe chromosome 3"
complement(1..32)
/note="rpl24-2 see c330 for full length CDS"
complement(709..1680)
/gene="SPCC320.14"
/note="synonym: SPCC330.15c"
complement(709..1680)
/gene="SPCC320.14"
/EC number="4.3.1.19"
/note="threonine dehydratase (predicted); involved in threonine catabolism; involved in branched chain amino acid biosynthesis; threonine dehydratase activity; localization unknown; similar to *S. cerevisiae* YKl218C PATH:MAP00260;
aspect=P; term=threonine catabolism; GOid=GO:0006567; evidence=ISS; db_xref=GOC:unpublished, with=SPTR:P05792; date=20030609
aspect=P; term=branched chain family amino acid biosynthesis; GOid=GO:0009082; evidence=ISS; db_xref=GOC:unpublished, with=SPTR:; date=20030609
aspect=F; term=threonine dehydratase activity; GOid=GO:0004794; evidence=ISS; db_xref=GOC:unpublished;

with=SPTR:; date=20030609
aspect=C; term=cellular component unknown;
GOid=GO:0008372; date=20030609"
/codon_start=1
/protein_id="CAA18316.1"
/db_xref="GI:2995376"
/db_xref="GOA:O59791"
/translation="MSDNLVLTPTDDVASERIKKIFANKTPVLTSTSTNKEFVAERVF
PCENFOKMGAFKFGALNALSQLNEAQKAGVLTFFSGNHAQAIASAKILGIPAKI
INPLDAPEAKVATKYGYQIMYDRYKDDREKAKEISEREGTLTIIPVDHPHVLG
QGTAAKELFEVPLDALFVCGGGLSGSALAKHAPNCEVGVVEAGNDGQGS
FRKGIHVITPTKTADGAQTHLGNFTSIIKEKVDVDDILTVSDEBILDLCKFYAARM
KIVVPTGCLSPAAARAMEKKNKRIIGIIISGNVDIERYAHFLSQ"
complement(1123..1518)
/gene="SPCC320.14"
/note="Pfam match to entry S_T_dehydratase ,
beta-synthase"
2454..2525
/gene="SPCTRNAGLU.09"
2454..2525
/gene="SPCTRNAGLU.09"
/product="tRNA-Glu"
/note="tRNA Glutamic acid
tRNA Glu anticodon CTC, Cove score 74.15"
3149..4216
/gene="ark1"
/note="synonyms: seel1, SPCC320.13c, SPCC330.16"
3149..4216
/gene="ark1"
/note="serine/threonine protein kinase; aurora homolog;
involved in spindle checkpoint; involved in spindle
formation (required) (PMID 11792803); involved in sister
chromatid separation (required) (PMID 11950927); involved
in the microtubule attachment response (required) (PMID
12676091); interacts physically and/or genetically with
the inner centromere protein Plcp (PMID 11950927);
interacts physically and/or genetically with the inner
centromere protein Birp (PMID 11950927); interacts
physically and/or genetically with survivin (PMID
11950927); involved in protein amino acid phosphorylation
(SER10 of Htt3p) (PMID 11792803); similar to *S. cerevisiae*
YPL209C
08.01.3003 trimmed to 2nd met PMID 12676091"
/codon_start=1
/protein_id="CAD88264.1"
/db_xref="GI:29786364"
/translation="MSDSKLADSLNCLSVSTSTTANPGRQQLLRLLAVSNQRVNVVS
LANGENKRTSNKSNKFNLSLRKTEEPFIAGVSSAGPQWREFHGMFEIGPLGKFGFR
VYLAKKKTFTVALKTLKSELVQSKIEKQVRRREIIOENLRHKNILRLYGHFHDDEK
RIVYLLEPAGRGELYOHLRAKRFSEEVASKYIFOMANALSYLHKHVIHRDIKPNKI
LLIGDGEIKLSDPGMSVHAPSNRRTLCGTLDYLPPEWVEGKHTKVDLSGLVLT
EFLVGAPPEDMSGHSATYKRIAKVDLKIPSPFPDPARDLISRLQLHNFKEKMSLSQV
MRHPWIVKYKDSWTRKSSSS"
3149..3151
/gene="ark1"
/note="initiator methionine, confirmed PMID 12676091"
3413..4168
/gene="ark1"
/note="Pfam match to entry pkinase PF00069, Eukaryotic
prot ein kinase domain"
3431..3502
/gene="ark1"
/note="PS00107 protein kinases ATP-binding region
signature"
3770..3808
/gene="ark1"
/note="PS00108 Serine/Threonine protein kinases
active-site signature"
complement(join(4707..4714,4781..5330))
/gene="SPCC320.12"
/note="synonym: SPCC330.17c"
complement(join(4707..4714,4781..5330))

misc_feature

gene

tRNA

gene

CDS

misc_feature

misc_feature

misc_feature

misc_feature

gene

CDS

STOVEVRDIIIDQFHEEDVTVFLLSTKAGGGINLACANVILVDCSYNPFDDLOAE	
DRAHVGQVRETVIRLIDNTIEYIKLANIKLADMSLSSDQKREERGERLUVQD	
MLDENKNTKPEITGNESDGEFKVSSNNKQDABETNTGVPLEGQPNKSVKTD	
LADGDKANIKTEKMSSETVEGDNRETKGNVQDNNAPSKSSTEPSNESVL	
SGHLDDTEASPVVSTIEKTKGDSVTEBQOSANIDGQLEKPEIEESKPDVLNOVS	
LSIEBKPNKESSEVDNNAKD"	
misc_feature	complement (8184..8435)
	/gene="SPCC1235.05c"
	/note="Match to helicase_C , Score 83.11"
	complement (8721..9638)
misc_feature	/gene="SPCC1235.05c"
	/note="Match to SNF2_N , Score 344.47"
Query Match 28.9%; Score 343.2; DB 8; Length 40880;	
Best Local Similarity 60.7%; Pred. No. 3.7e-81;	
Matches 580; Conservative 0; Mismatches 373; Indels 3; Gaps 1;	
Qy	202 AACACCTGGACCAAGAGAAATTAAAGCTATATATGACACACCACTCATGACCTGTATG 261
Db	
Qy	1434 AACAAATGACTCGTGAAGAAATTCAAAATAATTACGACACACCACTTATGATGATT 1493
Db	
Qy	262 CACTATGCTCAGGTGCAACACAGAAAGGTTCCAAACACCTTCAGAGGTTCAATTGTGCACT 321
Db	
Qy	1494 TTTAGAGCGCTTCATACATAGAAATTCACGACCCCAAGAAAGTTCACAATGCAT 1553
Db	
Qy	322 CTTATGAATATCAAACTGGTGTGTACCGAGGACTGTAAGTACTGTGCCCCAATCACAG 381
Db	
Qy	1554 TTACTTTCTATTAACAACAGGTGATACAGAAGATTGCMAATATTGTGCGCAATCTTCC 1613
Db	
Qy	382 CGTTACACACTGGTGTCAAGCTCAAGATCATCCAAAGTTGATGAGGTGATGAAGCT 441
Db	
Qy	1614 CGCTACAATACGGTGTCAAGGCCACTAAGCTCAAGAAATTTGACGAAGTTTGGAAAG 1673
Db	
Qy	442 GCAAAGGAGGCAAGCCAAATGGATCTACAAGGTTCTGTATGGGTGCTGCTTGAGAGAG 501
Db	
Qy	1674 GCTAAATGCAAGGCTTAAGGAGCACTAGATTTTGTATGGGAGTGCCTGGCGTAT 1733
Db	
Qy	502 ATGAAGTGAAGATCAAACTTGAAGAAAATCAAGAGAGATGATCACTGCTGTCTCATGAC 561
Db	
Qy	1734 TTGAATGTCGCAATCGCACCTTTAAAAATAATTTTGGAAATCATCAAGAGAGGTTGCGACG 1793
Db	
Qy	562 CTTGGAATGAGAGTGTGTACCTCGGATGTTGATTAAGACCAAGCCATGAATTG 621
Db	
Qy	1794 ATGATATGGAAGTTTGGCTAACCTCTGTATGTTAAATGAACAGCAAGCAAGCACTA 1853
Db	
Qy	622 AAAAGTGTGCTGCTGACCGGTACAAACCAATGATGATCTTCAAGAAACACTATCCA 681
Db	
Qy	1854 AAGGATGAGCTTGACGGCATACATCATATCTTGAATCTCTCGGATATATTACTCA 1913
Db	
Qy	682 AAGGTGATCTCCACAAGAAGCTTTGATGATAGATTGAAAACATTCAAAACGTTTCAAGGA 741
Db	
Qy	1914 AAGATTATCAGCACTAGGACATATGATGAACGGTTAAATACTATTGACAATCTTCGGAAA 1973
Db	
Qy	742 TCTGATTAAGGCATGACAGGTGTTCTTGTCTTGTGTGACCCCAAGAGACCGT 801
Db	
Qy	1974 GCTGGTTGAAGGTTTGTAGCGGTGTTATTTGGTCTAGGGGAGAAAAGCATGCTG 2033
Db	
Qy	802 GTATCTTCTCTACACCTGCGCAATGATGATCAGATCCAGATCTCTTCCATCAAC 861
Db	
Qy	2034 GTAGATTGATCTCGTTGAGAACTATGCAACACATCTCTGATCTGTACCTTTAAT 2093
Db	
Qy	862 AGACTGTTCCCAATCAAGGCGACCGCAATGTATGAAGAAGTTTAAGAAACAAGTGTAA 921
Db	
Qy	2094 CTTTTTGTCCGATTCCTGGCACCCCTGT ---AGGAGACGCTGTAAAGAGAGACTTCCG 2150
Db	
Qy	922 GTTATGATGATTGTGACACCATTTGCTACTTCAAGATTGGTCATGCCAAGCTTATTC 981
Db	
Qy	2151 ATTCACTCTTTCCTTCGCTCAATGCAACAGCTCGTATTTTGCATGCCAAGACTATCAT 2210
Db	
Qy	982 AGATTGGCTGAGGAAGATATACAATGAAGAGGAGCAAGAGGTGATGTCTTCATGGCT 1041
Db	
Qy	2211 CGTTTGACGCTGACGTAACGCTGCTCCGAGTCCGAGCAAGCACTTGTCTTTCATGGCT 2270
Db	

Qy	1042 GGTGTATGCACTTTCACAGGTAAGAAATGCTCAACAATGTGTACGGCTGGAT 1101
Db	
Qy	2271 GGTGTATGCACTTTCACCGGTGAAATAATGCTTACTCTCTGCTGTTCCTGGGAC 2330
Db	
Qy	1102 GAGGATAAGCCATGTTGGCTAAATGGGCTCTGAAACCAATGAGGAGTTTCAATA 1157
Db	
Qy	2331 TCAGACAGTCACTTTTACAAATGGGTTTGGAGGAATGCAAGGTTTGAATA 2386
Db	
RESULT 11	
Locus	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
FEATURES	
source	
misc_feature	
ORIGIN	
Query Match	

847 bp DNA linear STS 10-JAN-2001

T3 end of clone AT0AA003D06 of library AT0AA from strain CBS 4311

of Saccharomyces servazzii, sequence tagged site.

AL402619

AL402619.1 GI:12161819

STS

Saccharomyces servazzii

Saccharomyces servazzii

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

1 (bases 1 to 847)

Souciat J L, Aigle M, Artiguenave P, Blandin G, Bolotin-Fukuhara M, Bon E, Durrans P, Lepingle A, Liorente B, de-Montigny J, Dujon B, Durrans P, Lepingle A, Liorente B, Malpertuy A, Neuveglise C, Ozier-Kalogeropoulos O, Potier S, Saurin W, Tekala F, Toffano-Nioche C, Wesolowski-Louvel M, Wincker P, and Weissenbach J.

Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

FEBS Lett. 487 (1), 3-12 (2000)

1152876

2 (bases 1 to 847)

Casaregola S, Lepingle A, Bon E, Neuveglise C, Nguyen H, Artiguenave F, Wincker P, and Gaillardin C.

Genomic exploration of the hemiascomycetous yeasts: 7.

Saccharomyces servazzii

FEBS Lett. 487 (1), 47-51 (2000)

20584717

1152882

3 (bases 1 to 847)

Genoscope.

Direct Submission

Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. [E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]

This STS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Kluyveromyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

Location/Qualifiers

1..847

/organism="Saccharomyces servazzii"

/mol_type="genomic DNA"

/strain="CBS 4311"

/db_xref="taxon:27293"

/clone="AT0AA003D06"

/clone_lib="AT0AA"

/note="end : T3"

complement (<8..>847)

/note="similar to Saccharomyces cerevisiae ORF YGR286c [BI02 ; biotin synthetase]"

/evidence=not_experimental

28.0%; Score 332.2; DB 11; Length 847;

TITLE Stable macroscopic membranes formed by self-assembly of amphiphilic peptides and uses therefor
JOURNAL Patent: US 5670483-A 1 23-SEP-1997;
FEATURES Location/Qualifiers

source
1. 3083
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 27.6%; Score 328.2; DB 6; Length 3083;
Best Local Similarity 72.1%; Pred. No. 4.3e-77;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;
QY 553 GTCCATGACCTTGGAAATGGAGAGTTGTGTACCCCTGGGAATGGTTGATGAAGACCAAGCC 612
DB 6 GTGAATGATATGGGCTAGAAACGTTGTACTTTTAGGTATGTTGATCAAGATCAAGCA 65
QY 613 ACTGAATTGAAAAGTCTGGTTGACGGGTACCAACCATTAACATTAATGATCTTACAAGGAA 672
DB 66 AAGCAATTGAAAGATGACGGTTGACTGCATACCAACCATTAACATGACACATCTCCAGAGAA 125
QY 673 CACTATCCAAAGTGATCTCCCAAGAACGTTTGTATGATAGATTGAAAACATTTCAAAAC 732
DB 126 CACTATAGTAAAGTCTATCACCAGAACCTTACGACGACAGTTTACAGACCATCAAGAT 185
QY 733 GTTCAAGGATCTGGATTAAAGGATGACAGGTGGTATTTCTTGGTCTTGGTGAGACCCAA 792
DB 186 GTCCAAAGATCTGGAATAAAGCCTGTACCGTGGTATTTTGGTCTCGGTGAAGCGAA 245
QY 793 GAGGACCGTGTATCTTCTCTACACCTTTGGCCACAATGGATCAGCATCCAGAGTCTCTT 852
DB 246 GACGACCATATAGGATTCATCTACACATTTATCCAAATATGTCCTCATCTGAGTCCCTA 305
QY 853 CCAATCAACAGACTGGTCCCAATCAAGGCGACGCCAATGTATGAAGAAGTT-----903
DB 306 CCAATTAATAGACTAGTTGCTATCAAGGACTTCCAAATGGCTGAGGAATCTCCGATCCA 365
QY 904 AAGAAACAGCAAGTTGAAGTTGATGAGATTGTGAGAACCATTTGCTACTGCAAGATTGGTC 963
DB 366 AAGAGTAAAAAGTTGCAATTCGAGAAATTTTGAAGAACCATTTGCCACAGCGAGATAGTT 425
QY 964 ATGCCAAGCTTATATCAGATTGGCTGAGGAAGATATACAAATGAAGAGGCGAGAACAG 1023
DB 426 ATGCCAAAGGCCATTAATAGACTTGCCTGCTGTTATACAAATGAAGAAACAGAGCAA 485
QY 1024 GTGATGTCTTCCATGCTGTTTGAATGCCATCTTACAGGTAAGAAATGCTCAACA 1083
DB 486 TTTGCTGTTTCATGGCAGTTGTAAAGACTTGCCTGCTGTTATACAAATGAAGAAATGCTGACGACA 545
QY 1084 ATGTGTAACGGCTGGGATGAGGATTAAGCCATGTTGGCTAAATGGGCTGTAAGCAATG 1143
DB 546 ATATATAACGGTTGGGACGAGCAAGCAATGTTGGCTAAATGGGATGGCAACCTATG 605
QY 1144 GAGAGTTTCAATACAA 1160
DB 606 GAGCAATTTAAGTACGA 622

RESULT 14
AR306523
LOCUS AR306523 3083 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6548630.
ACCESSION AR306523
VERSION AR306523.1 GI:31696467
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3083)
AUTHORS Zhang, S., Lockshin, C., Rich, A. and Holmes, T.
TITLE Stable macroscopic membranes formed by self-assembly of amphiphilic peptides and uses therefor
JOURNAL Patent: US 6548630-A 1 15-APR-2003;

FEATURES Location/Qualifiers
source
1. 3083
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 27.6%; Score 328.2; DB 6; Length 3083;
Best Local Similarity 72.1%; Pred. No. 4.3e-77;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;
QY 553 GTCCATGACCTTGGAAATGGAGAGTTGTGTACCCCTGGGAATGGTTGATGAAGACCAAGCC 612
DB 6 GTGAATGATATGGGCTAGAAACGTTGTACTTTTAGGTATGTTGATCAAGATCAAGCA 65
QY 613 ACTGAATTGAAAAGTCTGGTTGACGGGTACCAACCATTAACATTAATGATCTTACAAGGAA 672
DB 66 AAGCAATTGAAAGATGACGGTTGACTGCATACCAACCATTAACATGACACATCTCCAGAGAA 125
QY 673 CACTATCCAAAGTGATCTCCCAAGAACGTTTGTATGATAGATTGAAAACATTTCAAAAC 732
DB 126 CACTATAGTAAAGTCTATCACCAGAACCTTACGACGACAGTTTACAGACCATCAAGAT 185
QY 733 GTTCAAGGATCTGGATTAAAGGATGACAGGTGGTATTTCTTGGTCTTGGTGAGACCCAA 792
DB 186 GTCCAAAGATCTGGAATAAAGCCTGTACCGTGGTATTTTGGTCTCGGTGAAGCGAA 245
QY 793 GAGGACCGTGTATCTTCTCTACACCTTTGGCCACAATGGATCAGCATCCAGAGTCTCTT 852
DB 246 GACGACCATATAGGATTCATCTACACATTTATCCAAATATGTCCTCATCTGAGTCCCTA 305
QY 853 CCAATCAACAGACTGGTCCCAATCAAGGCGACGCCAATGTATGAAGAAGTT-----903
DB 306 CCAATTAATAGACTAGTTGCTATCAAGGACTTCCAAATGGCTGAGGAATCTCCGATCCA 365
QY 904 AAGAAACAGCAAGTTGAAGTTGATGAGATTGTGAGAACCATTTGCTACTGCAAGATTGGTC 963
DB 366 AAGAGTAAAAAGTTGCAATTCGAGAAATTTTGAAGAACCATTTGCCACAGCGAGATAGTT 425
QY 964 ATGCCAAGCTTATATCAGATTGGCTGAGGAAGATATACAAATGAAGAGGCGAGAACAG 1023
DB 426 ATGCCAAAGGCCATTAATAGACTTGCCTGCTGTTATACAAATGAAGAAACAGAGCAA 485
QY 1024 GTGATGTCTTCCATGCTGTTTGAATGCCATCTTACAGGTAAGAAATGCTCAACA 1083
DB 486 TTTGCTGTTTCATGGCAGTTGTAAAGACTTGCCTGCTGTTATACAAATGAAGAAATGCTGACGACA 545
QY 1084 ATGTGTAACGGCTGGGATGAGGATTAAGCCATGTTGGCTAAATGGGCTGTAAGCAATG 1143
DB 546 ATATATAACGGTTGGGACGAGCAAGCAATGTTGGCTAAATGGGATGGCAACCTATG 605
QY 1144 GAGAGTTTCAATACAA 1160
DB 606 GAGCAATTTAAGTACGA 622

RESULT 15
SCZU01

LOCUS SCZU01 3083 bp DNA linear PLN 25-JUL-1993
DEFINITION S.cerevisiae ZU01 gene for zutotin.
ACCESSION X63612 S45839
VERSION X63612.1 GI:4836
KEYWORDS ZU01 gene; zutotin.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 3083)
AUTHORS Zhang, S., Lockshin, C., Herbert, A., Winter, E. and Rich, A.
TITLE Zutotin, a putative Z-DNA binding protein in Saccharomycetes
JOURNAL EMBO J. 11 (10), 3787-3796 (1992)
MEDLINE 93010971
PUBMED 1396572

```
REFERENCE 2 (bases 1 to 3083)
AUTHORS Zhang, S.
TITLE Direct Submission
JOURNAL Submitted (20-DEC-1991) S. Zhang, Massachusetts Inst of
Technology, 16-739 Dept of Biology, M I T, Cambridge MA 02139, USA
FEATURES
    source
        1..3083
            /organism="Saccharomyces cerevisiae"
            /mol_type="genomic DNA"
            /strain="208-12"
            /db_xref="taxon:4932"
            /chromosome="VII"
            /gene="ZUO1"
            1292..2593
            /gene="ZUO1"
            1292..2593
            /gene="ZUO1"
            /codon_start=1
            /product="put. zuotin"
            /protein_id="CAA45156.1"
            /db_xref="GI:4837"
            /db_xref="COA:P32527"
            /db_xref="SGD:S0003517"
            /db_xref="UniProt/Swiss-Prot:P32527"
            /translation="MFSPLTITSDITVEVNSSATKTPFVRRPVEPVGKFLQHAQRTL
            RNTWSEFERIEAEKNVKTVDSDNVDDELLFTLADLDELTHDARDMKTDADLYAAM
            GLSKLRPATESOIILKAHRKQVVKYHPDKOSAAAGSLDODGFFKIIOKAFETLTDNKK
            RAQYDCDFVADVPVPKGTDFDYFEAWGPVFEAEARFSKTPIPSLGNKDSKKKEVE
            QFYAHRFDSWRTFFELDSVDSDSNRDKRYIERKNKAARDKKKTADNARLVKLIV
            ERAVSEDPRIKMFKEBEKKEKREKREKREKREKREKREKREKREKREKREKREKREK
            SAKADKKKAAKAAKKNKRAIRNSKEADYFGDADKATIDEQVGLIVDSLNDEE
            LVSTADKIKANAAGAKEVLKESAKTIVDSKLPSSLISYFV"
            1914..2441
            /gene="ZUO1"
            /note="unnamed protein product; Protein sequence is in
            conflict with the conceptual translation; ORF within ZUO1"
            /codon_start=1
            /protein_id="CAA45157.1"
            /db_xref="GI:1326027"
            /translation="VTKILPRRLNNLSLSTDLTPGEPLSSWTKMSOMTLTETTSV
            TLKERTRPQTRRLRLITLWNSLLKELSVKIPVSKCKKKRRKREKENGKEKVPPE
            LKULRPRPKRLRLKLNKLPRLTPQKLTFRRLKLLRPPRRKTRTPSVTLURKLTT
            LVMLTRPRLTNKLIV"
            2746..2754
            polyA_signal
    ORIGIN
Query Match 27.6%; Score 328.2; DB 8; Length 3083;
Best Local Similarity 72.1%; Pred. No. 4.3e-77;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;
QY 553 GTCCATGACCTTGAATGGAGAGTTGTGTACACCTGGGAATGGTTGATAAGACCAAGCC 612
Db 6 GTGATGATATGGGGCTAGAAACGTGTACTTTAGGTATGGTTGATCAAGATCAAGCA 65
QY 613 ACTGAATTGAAAGTGTGGGTGGACGGGTGACCGGTACACCAATCAATGATCTTACAAGGAA 672
Db 66 AAGCAATTGAAGATGCAGGTTTGACTGTCATACCAACCAATCAACATCGACACTTCCAGAGAA 125
QY 673 CACTATCCAAAGGTGATCTCCACAAAGAGCTTTGATGATAGATGATAAATCAAAAC 732
Db 126 CACTATAGTAGGTGATCATCCACGAGAACTTACGACGACAGGTTACAGCCATCAAGAAT 185
QY 733 GTTCAAGGATCTGGATTAAGGATGACAGGTTGATTTCTTGGTCTTTGGTGAGACCAA 792
Db 186 GTCCAGAGATCTGGAATAAAGCCTGTACCGTGGTATTTTGGTCTCGGTGAAAGCGAA 245
QY 793 GAGGACCGTGTATCTTTCTCTACACCTTGGCCCAATGGATCAGCATCCAGAGTCTCTT 852
Db 246 GAGGACCATATAGGATTCATCTACATATATCCAAATATGCTCTCTCATCTGAGTCCCTA 305
QY 853 CCATCAACAGACTGGTCCCAATCAAGGGCAGCGCAATGATGAGAGATT----- 903
Db 306 CCAATTAAGACTAGTTGCTATCAAGGGGACTCCAAATGGCTGAGGAACCTTGGCGATCCA 365
```

Search completed: September 17, 2005, 17:41:52
Job time : 5416 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2005, 14:47:19 ; Search time 731 Seconds
(without alignments)
9620.592 Million cell updates/sec

Title: US-09-752-957F-1
Perfect score: 1188
Sequence: 1 atgcgtttattatgactgc.....agggtgcattcggtgcttga 1188

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870657 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1188	100.0	1188	9	ADA09588 Candida u
2	417.6	35.2	1125	12	ADP98553 C. albica
3	349.2	29.4	1659	6	ABA01978 Soybean b
4	349.2	29.4	1659	10	ADE76391 Ade76391 Soybean S
5	328.2	27.6	3083	2	AAT90969 Zuoatin ge
6	328.2	27.6	3083	2	AZ11218 Human zuo
7	328.2	27.6	3083	10	AAD59051 Yeast zuo
8	328.2	27.6	3083	12	ADO04472 Yeast Z-D
9	328.2	27.6	3083	13	ADS15330 Yeast DNA
10	311.8	26.2	1515	6	ABA01974 Maize bio
11	311.8	26.2	1515	10	ADE76383 Corn SID2
12	310.2	26.1	1313	3	AAC48072 Zea mays
13	310.2	26.1	1396	6	ABA01972 Ade76379
14	310.2	26.1	1396	6	ABA01975 Maize bio
15	310.2	26.1	1439	6	ABA01975 Maize bio
16	310.2	26.1	1439	10	ADE76385 Corn SID2
17	306.6	25.8	1152	6	ABA01967 Ade76369
18	306.6	25.8	1152	10	ADE76369 Corn bio
19	304	25.6	1137	6	ABZ12774 Arabidops
20	304	25.6	1137	10	ABV76106 Arabidops

21	304	25.6	1137	12	ADN74336 Adn74336 Thale cre
22	304	25.6	1340	3	AAC48445 Arabidops
23	304	25.6	1351	2	AAT34580 Biotin sy
24	304	25.6	1351	2	AAX01311 A. thalia
25	300.8	25.3	1342	3	AAC33135 Arabidops
26	295.6	24.9	1477	6	ABA01976 Prickly p
27	295.6	24.9	1477	10	ADE76387 Ade76387 Prickly p
28	294.4	24.8	1340	6	ABA01969 Soybean b
29	294.4	24.8	1340	10	ADE76373 Ade76373 Soybean b
30	292.8	24.6	1526	6	ABA01977 Soybean b
31	292.8	24.6	1526	10	ADE76389 Soybean S
32	292.6	24.6	1041	1	AAN91327 E.coli Bi
33	285.6	24.0	1248	13	ADR85519 Aspergill
34	285.6	24.0	1248	13	ADR84932 Aspergill
35	285.6	24.0	7248	13	ADR84345 Aspergill
36	260.8	22.0	1467	6	ABA01973 Maize bio
37	260.8	22.0	1467	10	ADE76381 Ade76381 Corn SID2
38	243.4	20.5	1032	6	ABA01979 Wheat bio
39	243.4	20.5	1032	10	ADE76393 Bread whe
40	237.6	20.0	987	8	ACA20541 Prokaryot
41	236	19.9	1017	9	ADA31411 DNA encod
42	228.6	19.2	957	8	ACA37222 Prokaryot
43	226.4	19.1	1038	10	ACF68915 Photornab
44	226.4	19.1	110000	10	ACF67367_17 Continuation (18 o
45	226.4	19.1	210710	10	ACF65380 Photornab

ALIGNMENTS

RESULT 1

ADA09588
ID ADA09588 standard; DNA; 1188 BP.

XX AC ADA09588;

XX DT 06-NOV-2003 (first entry)

XX DE Candida utilis biotin synthase gene.

XX KW Yeast; high biotin-productivity; edible yeast; biotin synthase gene;
XX RW feed additive; food additive; cosmetic; gene; ds.

XX OS Pichia jadinii.

XX PN US2003104584-A1.

XX PD 05-JUN-2003.

XX PF 02-JAN-2001; 2001US-00752957.

XX PR 07-OCT-2000; 2000TW-00120972.

XX PA (SHIU/) SHIUAN D.

XX PI Shiuian D;

XX DR WPI; 2003-567956/53.

XX PT New integrated plasmid, useful for preparing yeast with high biotin-productivity used as feed additives, food additives or cosmetics.

XX PS Claim 3; Page 3; 20pp; English.

XX CC The present invention relates to a yeast with high biotin-productivity and a method for its preparation. The yeast is transformed by an integrated plasmid, which includes the edible yeast (Candida utilis) biotin synthase gene, an assistant DNA sequence for the integration of the plasmid into a host genome, a promoter sequence, and a selection marker. The prepared yeast with high biotin-productivity or recovered biotin is useful as feed additives, food additives or cosmetics. The present sequence represents C. utilis biotin synthase gene.

SQ	Sequence	1188 BP; 376 A; 229 C; 295 G; 288 T; 0 U; 0 Other;	
	Query Match	100.0%; Score 1188; DB 9; Length 1188;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 1188; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGTCGTTTATATATGACTGCTATTAGTCGTCGATTCCTTCCACTCTTAGAGTAGCT	60
DB	1	ATGTCGTTTATATATGACTGCTATTAGTCGTCGATTCCTTCCACTCTTAGAGTAGCT	60
QY	61	TCTAGGCTACTCTGGCAACAGGTGCTACTGCTGCTCGGAGATCTTTGGAAGATGTTC	120
DB	61	TCTAGGCTACTCTGGCAACAGGTGCTACTGCTGCTCGGAGATCTTTGGAAGATGTTC	120
QY	121	ACGGAACAAATGCAAGAGTGGCTTACAGAGAGAGCAACCAATGGATATGCA	180
DB	121	ACGGAACAAATGCAAGAGTGGCTTACAGAGAGAGCAACCAATGGATATGCA	180
QY	181	TTGTCAGTGAAGACACCACTCAACACCTGGACCAAGAAATTAAGCTATATATGAC	240
DB	181	TTGTCAGTGAAGACACCACTCAACACCTGGACCAAGAAATTAAGCTATATATGAC	240
QY	241	ACACCACTATGACTGTGCTATGCTCAGTGCAACACAGAGTTCCAAACCT	300
DB	241	ACACCACTATGACTGTGCTATGCTCAGTGCAACACAGAGTTCCAAACCT	300
QY	301	TCAGAGTTCAATTTGTCACCTTATGAATATCAAACTGGTGTGTCACGAGGACTGT	360
DB	301	TCAGAGTTCAATTTGTCACCTTATGAATATCAAACTGGTGTGTCACGAGGACTGT	360
QY	361	AAGTACTGTGCCAATCACAGCGTTTACAACACTGGTGTCAAGGCTGAAGAATCATCCAA	420
DB	361	AAGTACTGTGCCAATCACAGCGTTTACAACACTGGTGTCAAGGCTGAAGAATCATCCAA	420
QY	421	GTTGATGAGTGATTTGAAGCTGCAAGAGGCAAGGCCAATGGATCTCAAGGTTCTGT	480
DB	421	GTTGATGAGTGATTTGAAGCTGCAAGAGGCAAGGCCAATGGATCTCAAGGTTCTGT	480
QY	481	ATGGGTCTCTTGGAGAGAGTCAAGGTAGAACTCAAACTTGAAGAAATCAAGAG	540
DB	481	ATGGGTCTCTTGGAGAGAGTCAAGGTAGAACTCAAACTTGAAGAAATCAAGAG	540
QY	541	ATGATCACTGCTGTCCATGACCTTGGAAATGAGAGTTGTGTCAACCTCGGAATGGTTGAT	600
DB	541	ATGATCACTGCTGTCCATGACCTTGGAAATGAGAGTTGTGTCAACCTCGGAATGGTTGAT	600
QY	601	AAAGACCAAGCCACTGAATTTGAAAAGTGTGGTTGACGGGTACCAACCAATCATTTGAT	660
DB	601	AAAGACCAAGCCACTGAATTTGAAAAGTGTGGTTGACGGGTACCAACCAATCATTTGAT	660
QY	661	ACTTACAGGAACACTATCCAAAGGTGATCTCCACAGAGAGCTTTGATGATAGTTGAA	720
DB	661	ACTTACAGGAACACTATCCAAAGGTGATCTCCACAGAGAGCTTTGATGATAGTTGAA	720
QY	721	ACATTCAAAAACGTTTCAAGGATCTGGATTAAGGCATGCAAGTGTATCTTTGGTCTT	780
DB	721	ACATTCAAAAACGTTTCAAGGATCTGGATTAAGGCATGCAAGTGTATCTTTGGTCTT	780
QY	781	GGTGAGACCAAGAGGACCGTGTATCTTCTTACACCTTGGCCCAATGGATCAGCAT	840
DB	781	GGTGAGACCAAGAGGACCGTGTATCTTCTTACACCTTGGCCCAATGGATCAGCAT	840
QY	841	CCAGAGTCTCTTCAATCAACAGACTGCTCCCAATCAAGGCGACGCCAATGTATGAAGAA	900
DB	841	CCAGAGTCTCTTCAATCAACAGACTGCTCCCAATCAAGGCGACGCCAATGTATGAAGAA	900
QY	901	GTTAAGAACCAAGCAAGTTGAAGTTGATGAGATTTGTCAGAACCAATTTGCTTCAAGATTG	960
DB	901	GTTAAGAACCAAGCAAGTTGAAGTTGATGAGATTTGTCAGAACCAATTTGCTTCAAGATTG	960
QY	961	GTCATGCCAAGCTTATATCAGATTGGCTGCGAGGAAGATATACAATGAAGAGGCGAGAA	1020
DB	961	GTCATGCCAAGCTTATATCAGATTGGCTGCGAGGAAGATATACAATGAAGAGGCGAGAA	1020

1021 CAGGTGATGTCTTTCATGGCTGTTGTAATGCCATCTTCACAGGTAAGAAAATGCTCACA 1080
1021 CAGGTGATGTCTTTCATGGCTGTTGTAATGCCATCTTCACAGGTAAGAAAATGCTCACA 1080
1081 ACAATGTGTAACGGCTGGGATGAGGATAAGCCATGTTGGCTAAATGGGGTCTGAACCA 1140
1081 ACAATGTGTAACGGCTGGGATGAGGATAAGCCATGTTGGCTAAATGGGGTCTGAACCA 1140
1141 ATGGAGAGTTTCAATATACAAACCAAGGAGGTTGCATTCGGTCTTGA 1188
1141 ATGGAGAGTTTCAATATACAAACCAAGGAGGTTGCATTCGGTCTTGA 1188

RESULT 2
ADP98553
ID ADP98553 standard; DNA; 1125 BP.
XX
AC ADP98553;
XX
AC
XX
DT 23-SEP-2004 (first entry)
XX
C. albicans specific gene, orf6.4105, DNA sequence.
DE
C. albicans specific gene, orf6.4105, DNA sequence.
XX
Diploid fungal cell; allele; gene disruption cassette;
KW promoter replacement fragment; antifungal; fungicide; gene therapy;
KW infection; Candida albicans; gene; ds.
XX
Candida albicans.
OS
WO2004056965-A2.
XX
PN
XX
PD 08-JUL-2004.
XX
PF 19-DEC-2003; 2003WO-US040618.
XX
PR 19-DEC-2002; 2002US-0434832P.
XX
PA (ELIT-) ELITRA PHARM INC.
PA (ELIT-) ELITRA CANADA LTD.
XX
PI Roemer T, Jjiang B, Boone C, Bussey H;
XX
WPI; 2004-500296/47.
DR P-PSDB; ADP98863.
XX
PT Constructing a strain of diploid fungal cells in which both alleles of a gene are modified comprises modifying the alleles of a gene in the fungal cells by recombination using a gene disruption cassette and a promoter replacement fragment.
XX
PS Claim 33; SEQ ID NO 6038; 163pp; English.
XX
CC The invention relates to a novel method for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The method comprises modifying the alleles of a gene in diploid fungal cells by recombination using a gene disruption cassette and a promoter replacement fragment. The invention further comprises: assembling a collection of diploid fungal cells each of which comprises modified alleles of a different gene; a strain of diploid fungal cells comprising modified alleles of a gene, where the first allele of the gene is inactivated by a gene disruption cassette comprising a nucleotide sequence encoding an expressible selectable marker; and the expression of the second allele of the gene is regulated by a heterologous promoter gene, and where the gene encodes the polypeptide mentioned above; a collection of diploid fungal strains comprising the diploid strains cited above, where substantially all the different genes that encode the above amino acid sequences are modified and are present in different diploid strains in the collection; a nucleic acid molecule microarray comprising nucleic acid molecules, where each nucleic acid molecule comprises a nucleotide sequence that is hybridizable to a target nucleotide sequence comprising any of the 310 nucleotide sequences listed in the

specification (ADP98516-ADP98825); identifying a gene that is essential to the survival or growth of a fungus, that contributes to the virulence and/or pathogenicity of a fungus, or that contributes to the resistance of a diploid fungus to an antifungal agent; identifying an antifungal agent that inhibits the growth of a diploid fungus, or a therapeutic agent for treatment of a mammalian disease; correlating changes in the levels of proteins or gene transcripts with the inhibition of growth or proliferation of a diploid fungal cell; a purified or isolated nucleic acid molecule comprising a nucleotide sequence encoding a gene product required for proliferation of *Candida albicans*, where the gene product consists of any of the above-mentioned amino acid sequences; a vector comprising a promoter operably linked to the nucleic acid molecule cited above; a host cell containing the vector; a purified or isolated polypeptide comprising any of the 61 amino acid sequences given in the specification (ADP96718-ADP96778); a fusion protein comprising a fragment of a first polypeptide fused to a second polypeptide, the fragment consisting of at least 6 consecutive residues of any of ADP98826-ADP99135; producing a polypeptide; identifying a compound which modulates the activity of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825; eliciting an immune response in an animal; a strain of *Candida albicans*, where a first allele of a gene comprising any of ADP98516-ADP98825 is inactive and a second allele of the gene is under the control of a heterologous promoter; identifying a compound or binding partner that binds to the polypeptide comprising any of ADP98826-ADP99135, or its fragment; identifying a compound having the ability to inhibit growth or proliferation of *Candida albicans*; inhibiting growth or proliferation of *Candida albicans* cells; manufacturing an antimycotic compound; treating an infection of a subject by *Candida albicans*; preventing or containing contamination of an object by *Candida albicans*, or for preventing or inhibiting formation of an object by *Candida albicans*, comprising *Candida albicans*; a pharmaceutical composition comprising a therapeutic amount of an agent which reduces the activity or level of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825 in a pharmaceutical carrier; an antibody preparation which binds the polypeptide; methods for evaluating a compound against a target gene product encoded by any of ADP98516-ADP98825; identifying an antimycotic compound; a computer or a computer readable medium that comprises at least one of the nucleotide sequences mentioned in the specification or at least one amino acid sequence selected from ADP98826-ADP99135; a method assisted by a computer for identifying a putatively essential gene of a fungus; and a protein array comprising proteins, where at least one protein comprises an amino acid sequence or a portion of an amino acid sequence selected from ADP98516-ADP98825. The novel methods and compositions have fungicide activity. The compositions may be used in gene therapy. The composition and methods are useful for drug screening purposes or for diagnosing, preventing or treating infections associated with *Candida albicans*. These may also be used for constructing strains useful for identification and validation of gene products as effective targets for therapeutic intervention, for identifying and validating gene products as effective targets for therapeutic intervention, and for collecting identified essential genes. This polynucleotide sequence represents a *Candida albicans* fungal specific gene of the invention.

NOTE: This sequence was downloaded from an electronic sequence listing provided on the WIPO website.

XX SQ Sequence 1125 BP; 362 A; 218 C; 242 G; 303 T; 0 U; 0 Other;

Query Match 35.2%; Score 417.6; DB 12; Length 1125;
 Best Local Similarity 67.0%; Pred. No. 6.7e-113;
 Matches 645; Conservative 0; Mismatches 299; Indels 18; Gaps 3;

QY 207 CTGGACCAAGAGAAATTAAGCTATATATGACACACCACTCATGGACTTGTGCACTA 266
 DB 159 CTGGACCAAGAGAAATTAAGCTATATATGACACACCACTCATGGACTTGTGCACTA 218

QY 267 TGCTCAGGTGCAACAGAGGTTCCAAACCTTCAGAGGTTCAATGTGCATCTTAT 326
 DB 219 CGCTCAATTGAACATCGTGAGTACCAAAAACCTGCTGAAGTCCCAATATGACTCTTGT 278

QY 327 GAATATCAAAACTCGTGGTGTACCGAGGAGCTGAAGTACTGTGCCCAATCACAGCGTTA 386
 DB 279 TAACATTAGCAGGGTGGATGACCGAAACTGTCTTACTGTCTCAATCATCGAAACA 338

QY 387 CAACACTGCTGCTCAAGGCTGAAGAAATCATCAAGTTGATGAGGTGATTGAAGCTGCAAA 446
 DB 339 TGATACTGGAGTTTCAAGCTGAAAAA---CTCGACTTGGATGCTGTGATGGACATTTGCTAG 395

QY 447 GGAGGCAAGGCCCAATGGATCTACAAAGTCTCTGTATGGTCTGCTGCTGGAGAGATGAA 506
 DB 396 AGATGCGAAGCAGAGGTGGTACTAGAGTTTGGCTTGGTCCGCTTGGAGAGATATGCA 455

QY 507 AGGTAGAAAGTCAAACTTTGAAGAAATCAAAAGAGATGATCACTGCTGCTCCATGAC---CT 563
 DB 456 AGGAAGAAGGCTCTCCCTTAGACAAAATCAGTAAAGTTGTTTAAACAAATCAACGACGAGTT 515

QY 564 TGGAAATGGAGAGTTGTGTCCCTCCCTGGGAATGGTTGATTAAGACCAAGCCACTGAATGAA 623
 DB 516 GAAGTTGGAAACTTTGTGTACTCTCGGGATGATCAACGAAAGAGCGGTGAAATTTTAA 575

QY 624 AAGTGTGGGTTGACGGGTACAACTTAACTTACAAAGGAACTACTATCCAAA 683
 DB 576 ACAGAAATGGTTGACTGCTTACCAACCAACATTTGACACTTCAAGAGAACATTTATCCAAA 635

QY 684 GGTGATCTCCAAGAAAGCTTTTGATGATAGATTGAAACCAATTCAAAAACGTTTCAAGGATC 743
 DB 636 CGTCGTCAACACCGGTATACAGCAAGAAAGATTAGAAACCATTTAAACACGTCCAAAAGC 695

QY 744 TGGATTAAAGCATGACAGGTGGTATTTCTTGGTCTTGGTGAGACCCCAAGAGACCGTGT 803
 DB 696 TGGTATCAAAAGCTTGTACCGGTGGTATTTCTCGGATTTGGGAGAAAACCGCACAGATCATGT 755

QY 804 ATCTTCTCTACACCTTGGCCCAATGGATCAGATCCAGAGTCTCTTCCAATCAACAG 863
 DB 756 TTGCTTTTGTGACACTTTGTCCCAATGAGTTCGCGACCCAGAAATCATTTGCCAATCAACAG 815

QY 864 ACTGGTCCCAATCAAGGGCAGCGCAATGTATGAA-----GAAAGTTAAGAAACA 911
 DB 816 ATTAGTTCCCAATCAGGGTACTCCAATTTGTGAAGTCTTGAATAATGCACCAAGGAAAG 875

QY 912 GCAAGTTGAAGTGTGATGAGATTGTTCAGAACCATTTGCTACTGCAAGATTGCTATGCCAAC 971
 DB 876 ACAATTGGAATTTGATGCTCTATTTTGAAGAACTATTTGCTACTGCTAGATTGATTGCTGTA 935

QY 972 GTCTATTATCAGATTGGCTCCAGGAAGATATACAAATCAAGAGGCGACAGGATGATGTG 1031
 DB 936 ATCCATCATTAGATTGGCCCGCGGTCTCATACCATCAAGGAAACACGAGCAGTCTTCTGTG 995

QY 1032 CTTTCATGGCTGGTTGTAATGCCATCTTCACAGGTGAAGAAATGCTCAACAAATGTGTAA 1091
 DB 996 TTTCATGAGTGGGTCAATGCTATTTTCCCGGTAAAGAGAAATGTTGACTCAATGTGTAA 1055

QY 1092 CGCTGGGATGAGATTAAGCCATGTTGGCTTAATGGGGTCTGAAACCAATGGAGAGTTT 1151
 DB 1056 TGGTTGGGATGAAGATATCGCCATGTTTGAAGAAATGGGGTTTGAACCAATGGAAAGTTT 1115

QY 1152 CA 1153
 DB 1116 CA 1117

RESULT 3

ABA01978

ID ABA01978 standard; cDNA; 1659 BP.

XX ABA01978;

DT 06-FEB-2002 (first entry)

DE Soybean biotin synthase coding sequence #4.

XX Soybean; biotin synthase; biotin biosynthesis; herbicide;
 KW plant development; transgenic plant; ss.

XX Glycine max.

XX Key

Location/Qualifiers

FT	CDS	47. .1300	
FT		/*tag= a	
FT		/product= "biotin synthase"	
XX			
XX	PN	US2001039042-A1.	
XX	XX	08-NOV-2001.	
XX	PD		
XX	PF	19-DEC-2000; 2000US-00740288.	
XX	PR	21-DEC-1999; 99US-0172929P.	
XX	XX	(ALLE/) ALLEN S M.	
PA	PA	(KINN/) KINNEY A J.	
PA	PA	(MIAO/) MIAO G.	
PA	PA	(OROZ/) OROZCO E M.	
XX	XX	Allen SM, Kinney AJ, Miao G, Orozco EM;	
XX	PI	WPI: 2002-040723/05.	
XX	DR	P-PSDB; AAMS1985.	
XX	DR		
XX	PT	New polypeptides, useful as targets for herbicide discovery, and as	
XX	PT	probes for genetic and physical mapping of genes of which they are part,	
XX	PT	or creating transgenic plants, comprises biotin synthase polypeptides and	
XX	PT	encoding polynucleotides.	
XX	PS	Claim 5; Page 35; 46pp; English.	
XX	CC	The present invention provides the protein and coding sequences of biotin	
XX	CC	synthase enzymes from barley, wheat, maize, soybean and the prickly	
XX	CC	poppy. These sequences can be used to produce transgenic plants which	
XX	CC	express different levels of the gene, or express it at different times in	
XX	CC	plant development. They can also be used as a target in the production of	
XX	CC	herbicides. The present sequence is a soybean biotin synthase cDNA	
XX	SQ	Sequence 1659 BP; 496 A; 334 C; 402 G; 427 T; 0 U; 0 Other;	
XX	Query Match	29.4%; Score 349.2; DB 6; Length 1659;	
XX	Best Local Similarity	60.6%; Pred. No. 1.6e-92;	
XX	Matches 595; Conservative	0; Mismatches 378; Indels 9; Gaps 1;	
QY	173	AAATGATGATGTCAGTGAAGACACAGTCAACACCTGGACCAAGAGAAATTAAGCTA	232
DB	234	AAGAGCACTCAAGCACTAAGTCCCGCAGCAATGGACAGAGAGAGAAATCAAGGGA	293
QY	233	TATATGACACACCACTCATGAGCTTTGATGCACTATGCTCAGGTGCAACACAGAGGTTCC	292
DB	294	TCTATGATAAGCCATTGATGAGTATGTTGGGGTGTGTTGACAGGAAATTC	353
QY	293	AACAACTTCAGAGGTTCAATTTGCACTCTTATGAATATCAAACTGGTGTGATACCG	352
DB	354	ATATACCTGGGGCTATTTCAGATGTTGATGTTGAAATCAAGACGGGTGTTGCTCGG	413
QY	353	AGGACTGTAACTACTGTGCCCAATCAACAGCTTACAACTGGTGTCAAGGCTGAAAGAA	412
DB	414	AGGATGTTCTTACTGCGCCCAATCATCCGCTACCAACCGGTCTCAAGCCTCCAAA	473
QY	413	TCATCCAAAGTTGATGAGGTGATGAAGCTGCAAGAGGAGGCAAGCCCAATGGATCTACAA	472
DB	474	TGGTCTCCGTGCAATCTGTCTCGCAGCGCCGCATCGCCAAAGACAACGGTAGTACAC	533
QY	473	GGTCTCTATGGGTGCTGTTGGAGAGAGATGAAGGTGAGAACTCAAACTTGAAGAAA	532
DB	534	GTTCCTGATGGGAGCGCGTGGCGCATATGCTGGACGGAACCAATCTCAAATG	593
QY	533	TCAAAGAGATGATCACTGCTGTCTCCATGACCTTGAATGGAGAGTTGTCAACCTCGGAA	592
DB	594	TCAAACAATGGTTAGCGAGATTCCGCGAATGGGTATGGAAATGATGTGTCAGCTTGTA	653
QY	593	TGGTTGATAAGAACCAAGCCACTGAATTTGAAAAGTGTGGTTGACGGCGTACACATA	652
DB	654	TGATTTGATGACAGCAAGCTCAGGAATCTCAAGAAAGCGGTCTCACGGCTTATAATCAT	713
QY	653	ACATTGATCTTACAAGGAACACTATCCAAAGGTGATCTCCAAAGAGCTTTGATGATA	712
DB	714	ATGTGATAGTCGAGGATTTCTATCCCAAGTTATCACGACCAAGACTTATGATGAGA	773
QY	713	GATTGAAAACATTTCAAAAGCTTCAAGGATCTCGATTAAGAGGATGACAGAGGTGATTC	772
DB	774	GATTGGATACCATTAAGAAATGTGAGAGAGCGCGGAATCAATGTTGTACGGGTGAATCC	833
QY	773	TTGGTCTTGGTGAGACCCCAAGAGCGCGTGTATCTTTCTCTACACCTTCGCCCAATGG	832
DB	834	TCGGATTAGGAGAAATAAGTCTGACCATATGAGACTTTTGGAGACGGTTGCTACGTTC	893
QY	833	ATCAGCATCCAGAGTCTCTTTCCCAATCAACAGACTGGTCCCAATCAAGGGCAACGCAATGT	892
DB	894	CTTCGCATCCGGAATCATTTCTCTGTGAACATGTTAGTGGCTATCAAGGAACACCACTGG	953
QY	893	ATGAAGAAGTTAAGAACCAAGCAAGTTGAGTTGATGAGATGTCAGAACCATTTGCTACTG	952
DB	954	AAGGA-----AACAGAGAGGTGGAATTTGAGAATATGTTGAGAATGTTTGCACGG	1004
QY	953	CRAGATTGGTCAATGCCCAACGCTATTATCAAGATTGGCTGCAGGAAGATATACAAATGAAG	1012
DB	1005	CTAGAATCGTCATGCCCTAAACCATCGTGGCTTTGGCAGCTGGAAGAGGAGAAATTGACG	1064
QY	1013	AGCAGAACACAGGTGATGCTTCATGGCTGGTGTGTTGTAATGCCATCTTTCACAGGTAAGAAA	1072
DB	1065	AGGAACAACAGGTCCTTATGTTTCATGGCGGAGCAATGCGTTTTCACAGGAGAAACAA	1124
QY	1073	TGCTCAACAATGTTGTAACGGCTGGATGAGGATGAAGCATGTTGGCTAAATGGGGTC	1132
DB	1125	TGTTAAACCAACACCAAGCGGTGGATGGGGTGTGCAATTCGCTCGTTCACAGATGGGGAT	1184
QY	1133	TGAACCAATGGAGAGTTTCAA 1154	
DB	1185	TAAGACCCATGGAAAGTTTCGA 1206	
XX	RESULT 4		
XX	ADE76391		
ID	ADE76391	standard; cDNA; 1659 BP.	
XX	AC	ADE76391;	
XX	AC		
DT	29-JAN-2004	(first entry)	
XX	DE	Soybean SID30 biotin synthase cDNA - SEQ ID 29.	
XX	DE		
XX	KW	biotin synthase; plant; breeding; seed; herbicide; ss; gene; soybean;	
XX	KW	SID30.	
XX	OS	Glycine max.	
XX	PN	US2003192073-A1.	
XX	PD	09-OCT-2003.	
XX	XX		
XX	XX	07-MAY-2003; 2003US-00431544.	
XX	XX		
PR	21-DEC-1999;	99US-0172929P.	
PR	19-DEC-2000;	2000US-00740288.	
XX	XX	(ALLE/) ALLEN S M.	
PA	PA	(KINN/) KINNEY A J.	
PA	PA	(MIAO/) MIAO G.	
PA	PA	(OROZ/) OROZCO E M.	
XX	XX	Allen SM, Kinney AJ, Miao G, Orozco EM;	
XX	PI	WPI: 2003-831831/77.	
XX	DR	P-PSDB; ADE76392.	
XX	PT	New isolated polynucleotide encoding plant biotin synthase, useful in	

erythropoietin, tissue-type plasminogen activator, synthetic haemoglobin and insulin. They can be used in applications, such as separation matrices (e.g. dialysis membranes). Collagen may be combined with the peptides to produce membranes for use as artificial skin. The MM may be used for making very thin, transparent fabric. Drugs which inhibit the self assembly of the peptides into filaments or filamentous membranes may be useful for treating Alzheimer's disease or scrapie infection. As they are resistant to proteolytic digestion and alkaline and acidic pH (such as stomach acid), drug delivery vehicles made of the MM could be taken orally. The charged residues and conformation of the MM promote cell adhesion and migration. The permeability of the MM also permits diffusion of small molecules, to the underside of cell monolayers, useful for tissue culture of differentiated cells and/or stratified cell layers. (Updated on 25-MAR-2003 to correct pf field.)

Sequence 3083 BP; 1026 A; 578 C; 627 G; 852 T; 0 U; 0 Other;

Query Match	27.6%;	Score 328.2;	DB 2;	Length 3083;
Best Local Similarity	72.1%;	Pred. NO. 3.8e-86;		
Matches 445;	Conservative	0;	Mismatches 163;	Indels 9;

553	Qy	GTCCATGACCTTGGAAATGGAGAGTTGTGTCAACCTGGGAAATGGTTGATAAAGACCAAGCC	612
6	Db	GTGAATGGATATGGGGCTAGAAAACGTTGTCTTCTTTAGGTATGGTTGATCAAGATCAAGCA	65
613	Qy	ACTCAATTTGAAAAGTCTGGGTGTGACGGCGTACAAACATAACATTGATCTATTACAAGAA	672
66	Db	AAGCAATTTGAAAAGATGCAAGTTTGACTGCATACAACCAATACATCGACACTTCCAGAGA	125
673	Qy	CACATATCCAAAGGTGATCTCCACAAGAAAGCTTTTGATGATAGATTGAAAAACATTCAAAAC	732
126	Db	CACATATAGTAAGGTCATCACCAAGAGAACTTACGACGACAGGTTACAGACCATCAAGAAT	185
733	Qy	GTTCGAAGATCTCGATTAAGGCGATGCACAGGTGGTATTCTTTGGTCTTTGGTGAGACCCAA	792
186	Db	GTCCAGAAATCTGGAATTAAGAGCCCTGTACCGGTGGTATTTTGGGTCTCGGTGAAGCGAA	245
793	Qy	GAGGACGCTGTATCTTTCTCTCACACCTTGGGCCAAATGGATCAGCATCCAGAGTCTCTT	852
246	Db	GACGACCATATAGGATTCATCTACACATATCCAATATGTCTCCTCATCTCTGAGTCCCTA	305
853	Qy	CCAATCAACAGACTGGTCCCAATCAAGGGCAGCCCAATGTTATGAAGAAGTT-----	903
306	Db	CCAATTAATAGACTAGTTGCTATCAAGGGGACTCCAATGGCTGAGGAACTTGGCGATCCA	365
904	Qy	AAGCAAGCAAGTTGAAGTTGATGAGATTGTCAGAACCATTTGCTACTGCAAGATTGTC	963
366	Db	AAGAGTAAAAAGTTGCAATTCGACGAAATTTTGAGAACCATTTGCCACAGCGAAGATGTT	425
964	Qy	ATGCCAACGCTCTATTATCAGATTGGCTGCAGGAAGATATACAATGAAAGAGCGAGAACAG	1023
426	Db	ATGCCAAGGCCAATTATAAGACTTGGCGCTGGTCTGTTATACAATGAAAGAAACAGAGCAA	485
1024	Qy	GTGATGTGCTTCATGGCTGGTTGTAAATGCCATCTTCAAGGTAAAGAAAATGCTCACACA	1083
486	Db	TTTGCTCTGTTTCATGGCAGGTTGTAAACAGTATCTTCAACGGTAAGAAAATGCTGACGCA	545
1084	Qy	ATGTGTAACGGCTGGGATGAGGATAAGCCATTTTGGCTAAATGGGGTCTGAAACCAATG	1143
546	Db	ATATATAACGGTTGGGACGAGACNAAGCAATGTTGGCTAAATGGGGATTGCACCTATG	605
1144	Qy	GAGAGTTTCAAAATCAA	1160
606	Db	GAGGCATTTAAGTACGA	622

RESULT 6

AAZ11218

ID AAZ11218 standard; DNA; 3083 BP.

XX

AC AAZ11218;

XX
DT 09-NOV-1999 (first entry)

XX		Human zuotin coding sequence.
DE		
XX		
KW		Membrane forming peptide; cell culturing; macroscopic membrane;
KW		amphiphilic peptide; slow-diffusion drug delivery system; cell growth;
KW		artificial skin; separation matrix; artificial tissue; scrapie infection;
KW		Alzheimer's disease; liver cirrhosis; kidney amyloidosis;
KW		protein conformational disease; human; zuotin; ZUO1; ss.
XX		
OS		Homo sapiens.
XX		
FH		Key Location/Qualifiers
CDS		1292..2593
FT		/tag= a
FT		/product= "zuotin"
XX		
PN		US9555343-A.
XX		
PD		21-SEP-1999.
XX		
PF		22-AUG-1994; 94US-00293284.
XX		
PR		28-DEC-1992; 92US-00973326.
XX		
PA		(MASI) MASSACHUSETTS INST TECHNOLOGY.
XX		
PI		Rich A, Dipersio CM, Lockshin C, Holmes T, Zhang S;
XX		
DR		WPI; 1999-539576/45.
DR		P-PSDB: AAY32954.

Sequence 3083 BP; 1026 A; 578 C; 627 G; 852 T; 0 U; 0 Other;

Query Match	27.6%;	Score 328.2;	DB 2;	Length 3083;
Best Local Similarity	72.1%;	Pred. No. 3.8e-86;		
Matches 445;	Conservative 0;	Mismatches 163;	Indels 9;	

Qy

553 GTCCATGACCTTTGGAA TGGAGATTGTGCACCCTCGGAATGGTTCATAAAGACCAAGCC 612
|| ||| | | | |
Dβ

6 GTGAATGATATGGGGCTAGA AACCTGTGTA CTTTAGTGATGGTTGATCAAGATCAAGCA 65
|| ||| | | | |

Db 246 GAGGACCATATAGGATTCATCTACACATTTATCCAAATATGTCTCTCATCTCGATCCCTA 305
Qy 853 CCATCAACAGACTGTGCCCAATCAAGGCGACGCCCAATGTATGAAGAAGTT----- 903
Db 306 CCAATTAAATAGACTAGTGTCTATCAAGGGAGCTCCAATGGCTGAGGAACCTTGGCGATCCA 365
Qy 904 AAGAACCAAGCAAGTTGAAGTTGATGAGATTTGTGAGAACCAATTTGCTACTGCAAGATTGGTC 963
Db 366 AAGAGTAAAGAAAGTTGCAATTCGACGAATTTTGAAGAACCAATTTGACACAGGAGAAATAGTT 425
Qy 964 ATGCCAACGCTTATTATCAGATGGCTGCGAGGAAGATATACATGAAGAGGAGCAACAG 1023
Db 426 ATGCCAAAGGCCATTAATGAAGCTTGGCGCTGGTGTATACATGAAGAAAGAACAGACAA 485
Qy 1024 GTGATGTCTTACGCTGGTGTGTAATGCCATCTTACAGCTGAAGAAATGCTCAACA 1083
Db 486 TTTGCTGTCTTATGCGAGGTTGTAACAGTATCTTACCGGTGAAGAAATGCTGACGACA 545
Qy 1084 ATGTGTAACGGCTGGGATAGGATAAAGCCATGTGTGGCTAAATGGGCTTGAACCAATG 1143
Db 546 ATATATAACGGTTGGGACGAAGCAAGGCAATGTGTGGCTAAATGGGATTTGCAACCTATG 605
Qy 1144 GAGAGTTTCAAAATACAA 1160
Db 606 GAGGCATTTAAGTACGA 622

RESULT 8

AD004472

XX AD004472 standard; DNA; 3083 BP.

AC AD004472;

XX 29-JUL-2004 (first entry)

XX Yeast Z-DNA binding protein (zuotin) DNA.

XX Macroscopic membrane; cell growth; EAK16; Z-DNA binding protein; zuotin;
KW medical product; suture; artificial skin; internal lining;
KW slow-diffusion drug delivery system; protein-type drug; erythropoietin;
KW haemoglobin; insulin; cell adhesion; cell migration; Alzheimer's disease;
KW scrapie infection; yeast; gene; ds.
XX Saccharomyces cerevisiae.

PH Key Location/Qualifiers
FT repeat_region 701..714
FT /*tag= b
FT /rpt_type= TANDEM
FT repeat_unit 701..702
FT /*tag= a
FT /note= "7 AT repeats"
FT repeat_region 850..859
FT /*tag= c
FT /rpt_type= TANDEM
FT repeat_region 950..958
FT /*tag= d
FT /rpt_type= TANDEM
FT repeat_region 1016..1031
FT /*tag= f
FT /rpt_type= TANDEM
FT repeat_unit 1016..1017
FT /*tag= e
FT repeat_region 1153..1162
FT /*tag= "8 AT repeats"
FT /rpt_type= g
FT /rpt_type= TANDEM
FT CDS 1292..1293
FT /*tag= h
FT /products= "Yeast Z-DNA binding protein (zuotin)"
FT misc_feature 2159..2206
FT /*tag= i
FT /note= "Non-purine tract"

XX US2004087013-A1.
PN 06-MAY-2004.
XX 17-MAR-2003; 2003US-00390472.
XX 28-DEC-1992; 92US-00973326.
PR 22-AUG-1994; 94US-00293284.
PR 26-MAR-1997; 97US-00824515.
XX (HOLM/) HOLMES T.
PA (ZHAN/) ZHANG S.
PA (RICH/) RICH A.
PA (DIPE/) DIPERSIO C M.
PA (LOCK/) LOCKSHIN C.
XX Holmes T, Zhang S, Rich A, Dipersio CM, Lockshin C;
PI WPI; 2004-356208/33.
XX P-PSDB; ADO04473.
XX Novel EAK16 protein incorporated into macroscopic membranes, useful in
PT biomaterial applications such as medical products, artificial skin or
PT internal linings, slow-diffusion drug delivery systems for in vitro cell
PT growth.
XX Claim 30; SEQ ID NO 1; 56pp; English.
XX The invention relates to a method for in vitro cell culture which
CC involves adding a macroscopic membrane that is formed by self-assembly of
CC amphiphilic peptide in an aqueous solution containing monovalent metal
CC cations to a cell culture medium comprising cells, thus forming a
CC membrane/culture mixture, maintaining the mixture under conditions
CC sufficient for cell growth. The invention also relates to EAK16 peptide
CC derived from yeast Z-DNA binding protein (zuotin). Zuotin incorporated
CC into the macroscopic membranes are useful in biomaterial applications
CC such as medical products (e.g., sutures), artificial skin or internal
CC linings, slow-diffusion drug delivery systems supports for in vitro cell
CC growth or culture and support for artificial tissue for in vivo use, as
CC slow-diffusion drug delivery vehicle for delivering protein-type drugs
CC e.g., erythropoietin, synthetic haemoglobin, insulin, etc., useful as
CC conductive biopolymer for culturing cell monolayers, for promoting cell
CC adhesion and migration, useful as experimental models for Alzheimer's
CC disease and scrapie infection. The present sequence is yeast zuotin DNA.
CC This sequence is used to illustrate the method of the invention.
XX Sequence 3083 BP; 1026 A; 578 C; 627 G; 852 T; 0 U; 0 Other;

Query Match 27.6%; Score 328.2; DB 12; Length 3083;
Best Local Similarity 72.1%; Pred. No. 3.8e-86;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;
Qy 553 GTCATGACCTTGGAAATGGAGAGTGTGTGCACCCCTGGGAATGGTTGATTAAGACCAAGCC 612
Db 6 GTCAATGATATGGGCTAGAAAACGTTGTTACTTTAGGTATGGTTGATCAAGATCAAGCA 65
Qy 613 ACTGAATTCAAAAGTCTGGGTTGACGGGTACACCAATTAACATTGATCTACTACAGGAA 672
Db 66 AAGCAATTCAAAAGATGCGAGGTTTGATGTCATACAAACCAATACATCGCACTTCCAGAGAA 125
Qy 673 CACTATCCAAAGGTGATCTTCCAAAGAGCTTTGATGATAGATTGAAAAACATTTCAAAAAC 732
Db 126 CACTATAGTAAGTTCATCACACGAGAACCTACGACGACAGGTTACAGACCATCAAGAT 185
Qy 733 GTTCAAGGATCTGGATTAAGGCGATGCACAGGTGGTATTTCTTGGTCTTGGTGAGACCAA 792
Db 186 GTCCAAGAAATCTGGAATAAAAGCCTGTACCGGTGGTATTTTGGGTCTCGGTGAAGCGAA 245
Qy 793 GAGGACCGGTATCTTTCTCTACACCTTGGCCACATGGATGCAGCATCCAGATCTCTT 852
Db 246 GACGACCATATAGGATTCATACATTATCAATATGTCCTCTCATCTGATGCCCTA 305

```

Qy 853 CCAATCAACAGACTGGTCCCAATCAAGGGCAGCCCAATGTATGAAGAAGTT----- 903
Db 306 CCAATTAATAGACTAGTTGCTATCAAGGGGACTCCCAATGGGTGAGGAACCTTGGCGATCCA 365
Qy 904 AAGAACAGCAAGTTGAGTTGATGAGATTGTCAGAACCAATGCTACTGCAAGATTGGTC 963
Dy 366 AAGAGTAAAGTTGCAATTCGACGAAATTTTGAGAACCAITGCCCACAGCGAGATAGTT 425
Qy 964 ATGCCAACGCTCTATTATCAGATTGGCTGCAGGAAGATATACAATGAAAGAGCGCAGAACAG 1023
Dy 426 ATGCCAAAGGCCATATATAGACTTGGCGTGGTCTGTATACATGAAAGAACAGAGCAA 485
Qy 1024 GTGATGCTTCATGCGTGGTGTGTAATGCCATCTTTCACAGGTAAGAAATGCTCACAACA 1083
Dy 486 TTTGCTGCTTTTCATGGCAGGTTGTAAACAGATATCTTCCCGGTGAGAAATGCTGACGACA 545
Qy 1084 ATGTGTACGCTGGGATGAGGATTAAGCCCATGTTGGCTAAATGGGCTGAAACCAATG 1143
Dy 546 ATATATAACGGTTGGGACGAGAACAGGCAATGTTGGCTAAATGGGGATTGCAACCTATG 605
Qy 1144 GAGAGTTTCAAAATACAA 1160
Dy 606 GAGGCATTAAGTACGA 622

RESULT 9
ADSL5330
ID ADS15330 standard; DNA; 3083 BP.
AC ADS15330;
XX
XX 16-DEC-2004 (first entry)
XX
XX Yeast DNA binding protein zootin DNA.
XX
XX cell culture; macroscopic membrane; amphiphilic peptide; biomaterial;
XX slow-diffusion drug delivery system; artificial skin; separation matrix;
XX Alzheimer's disease; scrapie; liver cirrhosis; kidney amyloidosis;
XX protein conformational disease; yeast; DNA binding protein; zootin; ds;
XX gene.
XX
XX Ascomycota.
XX
XX Key Location/Qualifiers
XX CDS 1292..2593
XX FT /*tag= a
XX FT /product= "Zootin"
XX
XX US6800481-B1.
XX
XX 05-OCT-2004.
XX
XX 26-MAR-1997; 97US-00824513.
XX
XX 28-DEC-1992; 92US-00973326.
XX 22-AUG-1994; 94US-00293284.
XX
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Holmes T, Zhang S, Rich A, Dipersio CM, Lockshin C;
XX WPI; 2004-707224/69.
XX
XX P-PSDB; ADS15331.
XX
XX Culturing cells, in vitro, by adding macroscopic membrane formed by self-
XX assembly of amphiphilic peptides and monovalent metal cations, to form
XX culture mixture, and maintaining culture mixture under conditions for
XX cell growth.
XX
XX Example 5; SEQ ID NO 1; 50pp; English.
XX
XX The invention describes a method of culturing (M1) cells, in vitro. The
XX cell culture involves adding a macroscopic membrane which is formed by

```

```

CC self-assembly of amphiphilic peptides in an aqueous solution containing
CC monovalent metal cations, where the peptides have alternating hydrophobic
CC and hydrophilic amino acids and are complementary and structurally
CC compatible, to a cell culture medium comprising cells, thus forming a
CC membrane/culture mixture, and maintaining the mixture under conditions
CC sufficient for cell growth. (M1) is useful for culturing a macroscopic
CC membrane utilised in biomaterial applications e.g., slow-diffusion drug
CC delivery systems, artificial skin or separation matrices, or as
CC experimental models for Alzheimer's disease and scrapie infectional
CC liver cirrhosis, kidney amyloidosis, or other protein conformational
CC diseases. (M1) enables in vitro culturing of macroscopic membrane that is
CC stable in aqueous solution, serum and ethanol, highly resistant to heat,
CC alkaline and acidic pH, chemical denaturation and proteolytic digestion,
CC and is non-cytotoxic. This sequence encodes yeast DNA binding protein
CC zootin from which amphiphilic peptides capable of forming membrane can be
CC isolated.
XX
XX Sequence 3083 BP; 1026 A; 578 C; 627 G; 852 T; 0 U; 0 Other;

```

```

Query Match 27.6%; Score 328.2; DB 13; Length 3083;
Best Local Similarity 72.1%; Pred. No. 3.8e-86;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;

```

```

Qy 553 GTCCATGACCTTGAATGGAGATTGTCACCCCTGGGAATGGTTGATAAGACCAAGCC 612
Dy 6 GTGAATGATATGGGGCTAGAAACGTGTGTTACTTTAGGTATGGTATCAAGATCAAGCA 65
Qy 613 ACTGAATTGAAAGTGTGGTTCACGGCTGACCAACCATACATTGATCTTACAAAGAA 672
Dy 66 AGCAATTTGAAGATGAGGTTTGCATGATCAACCAATACATCGACATTCGAGAA 125
Qy 673 CACTATCCAAAGGTGATCTCCCAAGAAAGCTTTGATGATAGATTGAAACAAATTCAAAAC 732
Dy 126 CACTATAGTAGTTCATCACCACGAGAACCTTACGACGACAGTTTACAGCCATCAAGAT 185
Qy 733 GTTCAAGATCTGAATTAAGGCATGACAGGTGGTATTCTTGGTCTTGGTGAGACCA 792
Dy 186 GTCCAAGAAATCTGAATAAAGCCCTGTACCGGTGGTATTTTGGGTCTCGGTGAAAGCGAA 245
Qy 793 GAGGACCGTGTATCTTTCTCTACACCTTGGCCCAATGGATCAGCATCCAGAGTCTCTT 852
Dy 246 GACGACCATATAGGATTCATCAATATCCAAATATGTCTCTCATCTCTGAGTCCCTA 305
Qy 853 CCAATCAACAGACTGGTCCCAATCAAGGGCAGCCCAATGTATGAAGAAGTT----- 903
Dy 306 CCAATTAATAGACTAGTTGCTATCAAGGGGACTCCAATGGCTGAGGAACCTTGGCGATCCA 365
Qy 904 AAGAACAGCAAGTTGAGTTGATGAGATTGTGAGAACCAATGCTACTGCAAGATTGGTC 963
Dy 366 AAGAGTAAAGTTGCAATTCGACGAAATTTTGAGAACCAITGCCCACAGCGAGATAGTT 425
Qy 964 ATGCCAACGCTCTATTATCAGATTGGCTGCAGGAAGATATACAATGAAAGAGCGCAGAACAG 1023
Dy 426 ATGCCAAAGGCCATTAAGACTTGGCGTGGTCTGTATACATGAAAGAACAGAGCAA 485
Qy 1024 GTGATGCTTCATGCGTGGTGTGTAATGCCATCTTTCACAGGTAAAGAAATGCTCACAACA 1083
Dy 486 TTTGCTGCTTTTCATGGCAGGTTGTAAACAGATATCTTCCCGGTGAGAAATGCTGACGACA 545
Qy 1084 ATGTGTAAACGGCTGGGATGAGGATTAAGCCCATGTTGGCTAAATGGGGTCTGAAACCAATG 1143
Dy 546 ATATATAACGGTTGGGACGAGAACAGGCAATGTTGGCTAAATGGGGATTGCAACCTATG 605
Qy 1144 GAGAGTTTCAAAATACAA 1160
Dy 606 GAGGCATTAAGTACGA 622

```

```

RESULT 10
ABAO1974
ID ABA01974 standard; cDNA; 1515 BP.
XX
XX ABA01974;
AC

```

XX 06-FEB-2002 (first entry)
 DT Maize biotin synthase coding sequence #5.
 DE
 XX Maize; biotin synthase; biotin biosynthesis; herbicide;
 KW plant development; clone cholic.pk009.j14; transgenic plant; ss.
 KW
 XX Zea mays.
 OS
 XX Key Location/Qualifiers
 FH 131..1264
 FT /*tag= a
 FT /product= "biotin synthase"
 FT
 PN US2001039042-A1.
 XX
 XX 08-NOV-2001.
 PD
 XX 19-DEC-2000; 2000US-00740288.
 XX
 XX 21-DEC-1999; 99US-0172929P.
 XX
 XX (ALLE/) ALLEN S M.
 PA (KINN/) KINNEY A J.
 PA (MIAO/) MIAO G. E.
 PA (OROZ/) OROZCO E M.
 XX
 XX Allen SM, Kinney AJ, Miao G, Orozco EM;
 PI
 XX WPI: 2002-040723/05.
 DR P-PSDB; RAMS1981.
 DR
 XX New polypeptides, useful as targets for herbicide discovery, and as
 PT probes for genetic and physical mapping of genes of which they are part,
 PT or creating transgenic plants, comprises biotin synthase polypeptides and
 PT encoding polynucleotides.
 XX
 XX Claim 5; Fig 2; 46pp; English.
 PS
 XX The present invention provides the protein and coding sequences of biotin
 CC synthase enzymes from barley, wheat, maize, soybean and the prickly
 CC poppy. These sequences can be used to produce transgenic plants which
 CC express different levels of the gene, or express it at different times in
 CC plant development. They can also be used as a target in the production of
 CC herbicides. The present sequence is a maize biotin synthase cDNA obtained
 CC from clone cholic.pk009.j14, which was derived from a maize embryo
 XX
 XX Sequence 1515 BP; 467 A; 368 G; 312 T; 0 U; 0 Other;
 SQ
 Query Match 26.2%; Score 311.8; DB 6; Length 1515;
 Best Local Similarity 59.1%; Fred. No. 2e-81;
 Matches 557; Conservative 0; Mismatches 377; Indels 9; Gaps 1;
 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTCATGGACTTGATG 261
 251 AACACTGGAGCGGCCGAGATCCAGCGCTCTACGACTCACGCTCTCGACTCCTC 310
 262 CACTATGCTCAGGTGCACACAGAGGTTCCAAACACCTTCAGAGGTTCAATGTGCACT 321
 311 TTTCAGGGGCTCAGGTCCACAGAAATGTCCATAAATTCAGAGAAATGCAGCAATGCACA 370
 322 CTTATGATATCAAACTGGTGTGTACCGGACTGTAGTACTGTGCGCAATCACAG 381
 371 CTTCTTTCAATCAAGACTGGTGGATGTCAGTGAAGATTGTTCTTACTGTCTCAGTCA 430
 382 CGTTACAACTGCTGTCAAGCTGAAAGAAATCATCCAAGTTGATGAGGTGATTGAAGCT 441
 431 AGATCAACACTGATTAAGAGCCCAAAATTAATGATGACAAAGATGCTGTCTTGAAGCA 490
 442 GCAAAGGAGGCAAGGCCAATGGATCTTACAAGGTTCTGATGGGTGCTGTGGAGAGAG 501
 491 GCAAAAAGGCAAAAGAGTCTCTGGAGGACCCCGTTTTCATGGGAGTGTGATGGAGAA 550

QY 502 ATGAAGGTAGAAAGTCAAACTTTGAAGAAAATCAAGAGATGATCACTGCTGTCCATGAC 561
 DB 551 ACCATTGGCAGGAAATCANAACTTCAACCAGATTCTTGAATATGTCAAGGAAATTAAGGGT 610
 QY 562 CTTGGAATGAGAGATTGTGTCACCTGGGAATGGTGTATTAAGACCAAGCACCCTGAATTG 621
 DB 611 ATGGCATGGAGGTCTGTGTCACACTAGGATGATAGAGAAACAACAAGCTGAAGAATC 670
 QY 622 AAAAGTCTGGTTGACGGGTACAACATAAATTGATATCTTACAAGGAAACATATCCA 681
 DB 671 AAGAGGCTGGACTTACAGCATATATCATTAACCTAGATATACAGAGAGTATATCCC 730
 QY 682 AAGGTGATCTCCACAAGAGCTTTTGATGATAGATTGAAAAATTCAAAAAAGTTCAAGGA 741
 DB 731 AACATTATTACCACAGATCATATGATGATAGTAGCTGCAGACTCTTTGAGCATGTCCGTGA 790
 QY 742 TCTGGATTAAAGGATGCACAGGTGGTATTTCTGGTCTTGGTGAACCAAGAGGACCGT 801
 DB 791 GCTGGAATAAGCATCTGCTCAGGTGGAATCATTTGGTCTTGGTGAAGCAGAGGAGCCGG 850
 QY 802 GTATCTTTCTCTACACCTTTGGCCACAATGATGATCAGCATCCAGAGTCTCTTCCAATCAAC 861
 DB 851 GTAGGTTGTTGTCATACCTTAGCTACCTTCCCTACACCCAGAGAGCGTTCTTATTAT 910
 QY 862 AGACTGGTCCCAATCAAGGGCAGCCCAATGTATGAAGAAGTTTAAGAAACAAGCAAGTTGAA 921
 DB 911 GCATTGGTTGCTGTAAAGGCACACCTCTTGAGG-----ACCAGAAGCCTGTAGAG 961
 QY 922 GTTGATGAGATTGTGAGAACCATTTGTACTGCAAGATTGCTGATGCCAAGCTCTATTATC 981
 DB 962 ATCTGGAAATGATCCGATGATCGCCACTGCTCGGATCAGATGCCAAGGCAATGGTG 1021
 QY 982 AGATTGGCTCAGAGAGATATACAATGAAGAGGAGCAGAGGATGATGCTTCCATCGCT 1041
 DB 1022 AGGCTTTTCAGCAGGCGGAGTACCGGTTCTCGATGCCAAGCAAGCGCTGTGCTTCTCGCT 1081
 QY 1042 GGTGTAATGCCATCTTTCAGAGGTAAGAAAATGCTCAACAATGTGTAAACGGCTGGGAT 1101
 DB 1082 GGGGCCAACTCCATCTTTGCGGGGAGAACTTCTCAACACCGCAACACGACTTTTGTAT 1141
 QY 1102 GAGATAAAGCCATGTTGGCTTAATGGGGTCTGAAACCAATGG 1144
 DB 1142 GCGGACCAGGCGATGTTCAAGATCCTTGGGCTGTATCCCCAAGG 1184
 RESULT 11
 ADE76383
 ID ADE76383 standard; cDNA; 1515 BP.
 XX AC ADE76383;
 XX AC ADE76383;
 XX 29-JAN-2004 (first entry)
 DT
 XX Corn SID22 biotin synthase cDNA - SEQ ID 21.
 DE
 XX biotin synthase; plant; breeding; seed; herbicide; ss; gene; maize; corn;
 KW SID22.
 KW
 XX Zea mays.
 OS
 XX
 XX US2003192073-A1.
 PN
 XX 09-OCT-2003.
 PD
 XX 07-MAY-2003; 2003US-00431544.
 XX
 XX 21-DEC-1999; 99US-0172929P.
 PR
 XX 19-DEC-2000; 2000US-00740288.
 PR
 XX (ALLE/) ALLEN S M.
 PA (KINN/) KINNEY A J.
 PA (MIAO/) MIAO G.
 PA


```
QY 322 CTTATGAATATCAAAAGCTGGTGGTGTGTACCGAGGACTGTAGTACTGTGCCATCAAG 381
Db 333 CTTCTTTCAATAAGACTGGTGGTGCAGGAGATTGTTTCATCTGCCACAGCTCTTCA 392
QY 382 CGTTACAACTGTGTCAAGGCTGAAGAAATCATCCAAGTTGATGAGGTGATTTGAAGCT 441
Db 393 AGATACAGTACCGGATTGAAGGCTGAAGAAATTAATGAAGAAAGATGCCGCTCTAGAAACA 452
QY 442 GCAAGAGGAGCAAGGCCAATGGATCTACAAGGTTCTGTATGGTGTGTCTGGTGGAGAG 501
Db 453 GCTAAAGAGCAAGGAGGCTGGAGCACCGGATTTTGCATGGAGCGCATGGAGAGAG 512
QY 502 ATGAAGAGTGAAGAGTCAAACTTGAAGAAATCAAGAGATGATCACTGTCTGCCATGAC 561
Db 513 ACAATTGGCAGGAAACAACTTCAACAGATTCTTGAATATGTCAAGGACATAAGAGGT 572
QY 562 CTTGGAATGAGAGTGTGTCACTGGGAAATGTTGATTAAGACCAAGCCACTGAATG 621
Db 573 ATGGGCATGGAGGTCTGTGTGACCCCTGGGCATGCTAGAGAAACAGCAAGCTGAAGAACTC 632
QY 622 ABAAGTCTGGTGTGACGGCTACAAACATTAACATTTGATCTTACAAGGAACACTATCCA 681
Db 633 AAGAGGCTGACTTACAGCTTATATCATTAACCTAGATACATCAAGAGATATATACCG 692
QY 682 AAGTGATCTCCACAAGAGCTTTTGATGATAGATTGAACATTTCAAAAAGCTTCAAGGA 741
Db 693 AACATTATTTCTACAAGATCGTATGATGATAGATTACAGACTCTTCAGCATGTCGGTGA 752
QY 742 TCTGGATTAAGGATGCAAGGTGGTATCTTGGTCTTGTGTGAGACCAAGAGGACCGT 801
Db 753 GCTGGAATAAGCGTCTGTCTAGGTTGGAATTAATGGTCTTGGAGAGCGGAGGAGACCGT 812
QY 802 GTATCTTCTCTACACTTGGCCACATGATGATCAGATCCAGAGTCTTCTTCCAAATCAAC 861
Db 813 GTAGGCTGTGTCATACACTGGCCACTTGGCCACATGTCACACACACCCAGAGAGTCTTCAAT 872
QY 862 AGACTGTCTCCAAATCAAGGCGACGCCAATGTATGAAGAAGTTAAGAAACAAGCAAGTTGAA 921
Db 873 GCATTGATGTGTCAAGGCAAGCGCTCTTCAGG-----ATCAGAGCCTGTAGAG 923
QY 922 GTGATGAGATTGTCAAGCAATTTGTCTACTGCAAGATTGTCATGCAAGCTCTATTATC 981
Db 924 ATATGGGAAATGATCCGATGATTGCCAGCGCTCGGATTTGTATGTCGCAAGGCAATGGT 983
QY 982 AGATTGCTGCGAGAGATATACATGAAGAGGAGAGAGGAGTGTCTTCTATGCT 1041
Db 984 AGACTTTCGGCAGGCGAGTACGGTTCTCCATGCCAGAGCAAGCTCTCTGCTTCTTCTGCT 1043
QY 1042 GGTGTAATGCTATCTTCAAGGTAAGAAATGCTCAACAATGTGTAAAGGCTGGGAT 1101
Db 1044 GGGCCCAACTCGATCTTCGCGGTGAAGAGCTCTGACAACTGCAAAACAAGCAAGCTTTGAT 1103
QY 1102 GAGGATAAGCCATGTTGGCTAAATGGGCTGTGAACCAATGG 1144
Db 1104 GCGGACCAGGCAATGTTCAAGATCCTTGGCCTGATTTCCCAAG 1146
```

RESULT 14

ADE76379
ID ADE76379 standard; cDNA; 1396 BP.

XX AC ADE76379;

XX AC ADE76379;

XX AC ADE76379;

XX AC ADE76379;

XX AC ADE76379;

XX AC ADE76379;

XX AC ADE76379;

XX AC ADE76379;

XX 09-OCT-2003.
PD XX
PF XX
XX 07-MAY-2003; 2003US-00431544.
XX 21-DEC-1999; 99US-0172929P.
PR 19-DEC-2000; 2000US-00740288.
XX (ALLE/) ALLEN S M.
PA (KINN/) KINNEY A J.
PA (MIAO/) MIAO G.
PA (OROZ/) OROZCO B M.

XX Allen SM, Kinney AJ, Miao G, Orozco EM;
XX WPI; 2003-831831/77.
XX P-PSDB; ADE76380.

New isolated polynucleotide encoding plant biotin synthase, useful in plant molecular biology and plant breeding purposes, particularly in the production of altered levels of biotin synthase in plants and seeds.
Claim 5; SEQ ID NO 17; 47pp; English.

The invention relates to a novel isolated polynucleotide encoding a first polypeptide or a third polypeptide where the first or third polypeptide is a biotin synthase. The methods and compositions of the invention may be useful in plant molecular biology and breeding, particularly in the production of altered levels of biotin synthase in plants and seeds. The polypeptides may be used as a target to facilitate the design and/or identification of inhibitors of these enzymes that may be useful as for herbicides. Furthermore, the polynucleotides may be used as probes for genetical and physically mapping the genes of which they are a part. The current sequence is that of the barley biotin synthase-related cDNA of the invention.

XX SQ Sequence 1396 BP; 387 A; 351 C; 356 G; 302 T; 0 U; 0 Other;

Query Match 26.1%; Score 310.2; DB 10; Length 1396;
Best Local Similarity 59.0%; Pred. No. 5.6e-81;
Matches 556; Conservative 0; Mismatches 378; Indels 9; Gaps 1;

QY 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACACCTCATGGACTTGATG 261
Db 213 AACGACTGGACCGCCCGAGATCCAGGCTATAGGCTCCCGCTCCGACCTCTC 272
QY 262 CATATGCTCAGGTGCAACACAGAGGTCCCAACACCTTCAGAGTTCAATTTGTCAC 321
Db 273 TTCACGGGGCTCAAGTCCATAGGAATGCCATAAATTTAGAGAGTGCACCAATGCAC 332
QY 322 CTTATGAATATCAAAACTGGTGGTGTGTACCGAGGACTGTAAAGTACTGTGCCCAATCA 381
Db 333 CTTCTTTCAATAAGACTGGTGGTGCAGCGAAGATTGTTTCATCTGCCACAGCTTCA 392
QY 382 CGTTACAACTGTGTCAAGGCTGAAGAAATCATCCAAGTTGATGAGGTGATTTGAAGCT 441
Db 393 AGATACAGTACCGGATTGAAGGCTGAAAAATTAATGAAGAAAGATGCCGCTCTAGAAACA 452
QY 442 GCAAGGAGGCAAGGCCAATGGATCTACAAGGTTCTGTATGGTGTGTCTGGAGAGAG 501
Db 453 GCTAAAGAGCAAGGAGGCTGGAGCACCGGATTTTGCATGGAGCGCATGGAGAGAG 512
QY 502 ATGAAGAGTGAAGAGTCAAACTTGAAGAAATCAAGAGATGATCACTGTCTGCCATGAC 561
Db 513 ACAATTGGCAGGAAACAACTTCAACAGATTCTTGAATATGTCAAGGACATAAGAGGT 572
QY 562 CTTGGAATGAGAGTGTGTCACTGGGAAATGTTGATTAAGACCAAGCCACTGAATG 621
Db 573 ATGGCATGGAGGTCTGTGTGACCCCTGGGCATGCTAGAGAAACAGCAAGCTGAAGAACTC 632
QY 622 ABAAGTCTGGTGTGACGGCTACAAACATTAACATTTGATCTTACAAGGAACACTATCCA 681
Db 633 AAGAGGCTGACTTACAGCTTATATCATTAACCTAGATACATCAAGAGATATATACCG 692

Qy	682	AAGGTGATCTCCACAGAAGCTTTGATGATAGATTGAAAAA	741
Db	693	AACATATTTTACAAAGATCGTATGATAGATTACAGACTCTT	752
Qy	742	TCTGGATTTAAAGGCAATGACAGGTGGTATTCTTGCTTTGGT	801
Db	753	GCTGGAAATTAAGCGTCTGCTCAGGTGGAAATTTGGTCTTGG	812
Qy	802	GTAATCTTTCCTCTACACCTTGGCCACAATGGATCAGCATCAG	861
Db	813	GTAGGGCTGTGGCATACACTGGCCACTTTGCCAACACACCCAG	872
Qy	862	AGACTGGTCCCAATCAAGGGCAGCGCAATGTATGAAGAAGTTA	921
Db	873	GCATTGATTGTGTCAAAAGGCACGCTCTTCAGG-----ATC	923
Qy	922	GTTGATGAGATTGTCAGAACCATTTGCTACTCCAGATTGGTCA	981
Db	924	ATATGGGAAATGATCCGCATGATTGCCACGCGCTCGGATTGAT	983
Qy	982	AGATTGGCTGCAGGAAGATATACAATGAAGAAGGCAGAACAGG	1041
Db	984	AGACTTTCGGCAGGCGGAGTACGGTTCTCCATGCCAGAGCAAG	1043
Qy	1042	GGTTGTAATGCACATCTTCACAGGTAAGAAATGCTCAACAATG	1101
Db	1044	GGGGCCAACTCGATCTTCGCCGGTGAAGAAGCTCTTCGACACT	1103
Qy	1102	GAGGATAAAGCCATGTTGGCTAAATGGGGTCTGAAACCAATGG	1144
Db	1104	GCGACACAGGAATGTTCAAGATCCTTGGCCTGATTTCCCAAGG	1146

RESULT 15

	ABA01975	ID ABA01975 standard; cDNA; 1439 BP.
XX	AC	
XX	ABAO1975;	
XX		
XX	06-FEB-2002	(first entry)
DE	XX	Maize biotin synthase coding sequence #6.
DE	XX	
XX		
KW	Maizer; biotin synthase; biotin biosynthesis; herbicide;	
KW	plant development; clone cca.pk0012.g11:fis; transgenic plant; ss.	
XX		
XX	Zea mays.	
OS		
XX		
Key	Location/Qualifiers	
FH	CDS	115..1248
FT	/tag= a	
FT	/product= "biotin synthase"	
XX		
US	U02001039042-A1.	
XX		
PB	08-NOV-2001.	
PD		
XX		
PF	19-DEC-2000; 2000US-00740288.	
PR		
XX	21-DEC-1999; 99US-0172929P.	
XX		
PA	(ALLE/) ALLEN S M.	
PI	(KINN/) KINNEY A J.	
PA	(MIAO/) MIAO G.	
PA	(OROZ/) OROZCO E M.	
XX		
PI	Allen SM, Kinney AJ, Miao G, Orozco EM;	
XX		
WI	WPI; 2002-040723/05.	
DR	P-PSDB; AAMS1982.	
XX		
PT	New polypeptides, useful as targets for herbicide discovery, and as	

PT probes for genetic and physical mapping of genes of which they are part,
PT or creating transgenic plants, comprises biotin synthase polypeptides and
PT encoding polynucleotides.

Query Match	26.1%	Score 310.2	DB 6	Length 1439
Best Local Similarity	59.0%	Pred. No. 5.7e-81		
Matches 556	Conservative 0	Mismatches 378	Indels 9	Gaps 1
Qy	202	AACACCTGGACCAAGAAGAAATTTAAAGCTATATATGACACACACACACATCATGGACTTGTGATG	261	
Db	235	AACACCTGGAGCCGGCCCGAGATCCAGCCGCTCTACGACTCACCCTCTCTCGACTCTCTC	294	
Qy	262	CACATATGCTCAGGTGCAACACAGAAAGGTTTCCAACAACCTTTACAGAGGTTCAATTTGTGCACT	321	
Db	295	TTTCACGGGCTCAGGTCCACAGAAATGTCCATAAATTCAGAGAAGTGCAGCAATGCACA	354	
Qy	322	CTTATGAATATCAAACTGGTGTGTACCGAGGACTGTAAGTACTGTGCCAATCAAG	381	
Db	355	CTTCTTTTCAATCAAGATCTGTGGATGCAGTGAAGATTGTTCTTACTGTCTCAGTCATCA	411	
Qy	382	CGTTACAACTCGTGTCAAGGGTGAAGAATCATCCAAGTTGATGAGGTGATTTGAAGCT	441	
Db	415	AGATACAACTGGATTGAAGGCCCAAAATTTGATGAACAATATGCTGCTTGGAGCA	474	
Qy	442	GCAAGAGGCAAGGCCAATGGATCTAACAGGTTCTGTATGGGTGCTGCTTGGAGAGAG	501	
Db	475	GCAAAAAGGCAAAAGAGTCTGGAGCACCCTGTTTTCATGGAGGCTCATGGAGAGAA	534	
Qy	502	ATGAAGGTGAGAAGTCAAACCTTTGAAGAAATCAAAGAGATGATCACTGCTGTCCATGAC	561	
Db	535	ACCATTTGGCAGGAAATCAAACTTCAACCAAGTCTTTGAATATGTCAGGAAATAAGGGGT	594	
Qy	562	CTTGGAAATGAGAGTTGTGTGTCACCTGGGGAATGGTTGATTAAGACCAGCCACTGAATTG	621	
Db	595	ATGGCATGAGGCTCTGTTGCACACTAGGCATGATAGAGAGAAACAACAGCTGAAGAACTC	654	
Qy	622	AAAAGTCTGGGTTGACGGGTACACCAATTAACATTTGATACTTACAGGGAACACTATCCA	681	
Db	655	AAGAAGCTGGACTTACAGCATATATCATPAACCTAGATACATCAAGAGAGTATTATCCC	714	
Qy	682	AAGGTGATCTCCACAAGAAGCTTTGATGATAGATTGAAAAATTTCAAAAACGTTTCAAGGA	741	
Db	715	AACATTATTACCACAAGATCATATGATGATAGACTGCAGACTCTTGAGCATGTCCCGTAA	774	
Qy	742	TCGTGGATTAAGGCATGCACAGTGGTATTCTTTGGTCTTGGTAGACCCCAAGGAGCCGT	801	
Db	775	GCTGGAATAAGCATCTGCTCAGGTGGAATCATTTGGTCTTGGTGAAGCAGAGGAGGACCG	834	
Qy	802	GTATCTTTCTCTACACACTTGGGCCAATGGGATCAGCATCCAGAGTCTCTTCCAATCAAC	861	
Db	835	GTAGGGTTGTTGCATACCTTAGTACTCTTGCTTACACACCAGAGAGCGTTCCTATTAT	894	
Qy	862	AGACTGGTCCCAATCAAGGGCAGCCCAATGTATGAAGAAGTTAAGAACAAAGCAAGTTGAA	921	
Db	895	GCATTGGTTGCTGTA AAAAGGCACACCTCTTTGAGG-----ACCAGAAGCCCTGTAGAG	945	
Qy	922	GTTGATGAGATTCTCAGAACCATTTGCTACTGCAAGATTGGTTCATGCCACAGCTTATTATC	981	
Db	946	ATCTGGGAATGATCCGATGATCGCCACTGCTCTCGGATCAAGTATGCAAGGCAATGGTG	1005	
Qy	982	AGATTGGCTGCAGGAAGATATACAAATGAAGAAGGCGAGAACAGGTGATGTGCTTCATGGCT	1041	

Db	1006	AGGCTTTACGACGCCGAGTACGGTTCTCGATGCCAGAACAAAGCGCTGTGCTTCTCGCT	1065
Qy	1042	GGTTGTAATGCCATCTTCACAGGTAAAGAAATGCTCACAACAATGTGTACGGCTGGAT	1101
Db	1066	GGGGCCAACTCCATCTTTGCCGGGAGAACTTCTCACAACCCGCAACACGACTTTGAT	1125
Qy	1102	GAGGATAAAGCCATGTTGGCTAAATGGGGTCTGAAACCAATGG	1144
Db	1126	GCGGACCGCGATGTTCAAGATCCTTGGCTGTATCCCCAAGG	1168

Search completed: September 17, 2005, 17:54:06
Job time : 737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2005, 16:11:49 ; Search time 236 Seconds
(without alignments)

8236.851 Million cell updates/sec

Title: US-09-752-957F-1

Perfect score: 1188

Sequence: 1 atgctgtttatattgactgc.....agggtgattcgggtgttga 1188

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	414.4	34.9	1263	4	US-09-248-796A-4445
2	328.2	27.6	3083	1	US-08-346-849-1
3	328.2	27.6	3083	2	US-08-293-284A-1
4	328.2	27.6	3083	4	US-08-898-300-1
5	328.2	27.6	3083	4	US-08-824-513-1
6	304	25.6	1351	2	US-08-401-068-13
7	304	25.6	1351	2	US-08-846-338-13
8	236	19.9	1017	4	US-09-328-352-2698
9	230.4	19.4	954	4	US-09-902-540-7391
10	230.4	19.4	4183	4	US-09-902-540-709
11	211.6	17.8	1830121	4	US-09-557-884-1
12	211.6	17.8	1830121	4	US-09-643-990A-1
13	208.2	17.5	1230025	4	US-09-198-452A-1
14	208.2	17.5	1230230	4	US-09-438-185A-1
15	205.2	17.3	1420	3	US-09-180-109A-23
16	203.6	17.1	1336	3	US-09-180-109A-28
17	200	16.8	640681	4	US-09-790-988-1
18	198.8	16.7	1053	4	US-09-543-681A-1175
19	195.2	16.4	1182	4	US-09-252-991A-3220
20	195.2	16.4	3150	4	US-09-252-991A-3348
21	193.8	16.3	1041	2	US-08-401-068-7
22	193.8	16.3	1041	2	US-08-846-338-7
23	192.4	16.2	5872	3	US-08-411-768B-1
24	192.4	16.2	5872	3	US-08-411-768B-6
25	191	16.1	1358	3	US-09-180-109A-25
26	186.4	15.7	3501	4	US-09-809-665A-70
27	184.2	15.5	34001	4	US-09-596-002-18

ALIGNMENTS

RESULT 1

US-09-248-796A-4445
; Sequence 4445, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICP
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4445
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Candida albicans
; US-09-248-796A-4445

Query Match 34.9%; Score 414.4; DB 4; Length 1263;
Best Local Similarity 66.8%; Pred. No. 3.4e-122;
Matches 643; Conservative 0; Mismatches 301; Indels 18; Gaps 3;

Qy	207	CTGGACCAAGAAATTAAAGCTATATATGACACACCACTCATGGACTTGATGCACTA	266
Db	297	CTGGACCAAGATGAATTAAGCAATTTACATGCTCCATGATGATTTGATTTCAA	356
Qy	267	TGCTCAGGTGCAACACAGAGGTTCCAAACACCTTCAGAGGTTCAATTTGSCATCTTAT	326
Db	357	CGTCAATTGAAACATCGTGAGTACCAAAACCTCTGGAAGTCCATTAATGACTCTTGT	416
Qy	327	GAATATCAAACTGGTGGTTGTACCGAGGCTTAAGTACTGTGCCCAATCACAGGTTA	386
Db	417	TAACATTAAGCAGGGTGGATGTACCGAAACCTGTTCTTACTGTTCTCAATCATCGAACA	476
Qy	387	CAACACTGGTGTCAAGGCTGAAAGATCATCCAAAGTTGATGAGTGTGATGAAGTCGAAA	446
Db	477	TGATACTGGAGTTCAAGCTGAAAAA---CTTGACTGGATGCTGTGATGACATTTGCTAG	533
Qy	447	GGAGGCAAGGGCAATGGATCTACAAAGTTCTGTATGGGTGCTGCTTGGAGAGATGAA	506
Db	534	AGATGCAAGACGAGGTGGTACTAGAGTTTGCTTGGTTGGCTTGGAGATATGCA	593
Qy	507	AGTGAAGAGTCAAACTTGAAGAAATCAAGAGATGATCAGTGTGTCATGAC---CT	563
Db	594	AGGAAGAAAGTCTGCCTTAGACAAATCAGTAAAGTTGTTAAACAAATCAACGACGATT	653

Sequence 5954, Ap
Sequence 1834, Ap
Sequence 338, App
Sequence 338, App
Sequence 7, Appli
Sequence 7, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 634, App
Sequence 539, App
Sequence 3541, Ap
Sequence 3895, Ap
Sequence 2561, Ap
Sequence 3934, Ap
Sequence 14, Appli
Sequence 1, Appli

564	Qy	TGGAATGGAGAGTTGTGTCAACCTGGGAATGGTTGATTAAGACAACGCCACTGAATTGAA	623
565			
654	Db	GAAGTTGGAAACTTGTGTACTCTCGGGATGATCAACGAAAGCAGCGCTGAAATTTTTAAA	713
655			
624	Qy	AAGTCTGGGTTGACGGCGTACAAACCATAACATTGATACTTTACAAGGAAACACTATCCAAA	683
625			
714	Db	ACAGAATGGTTTGACTGGCTTACAACCAACAATTGACACTTCAAGAGAACATATATCCAAA	773
715			
684	Qy	GGTGATCTCCACAAGAAAGCTTTGATGATAGATNTGAAACACATTCAAAAACGTTCAAGGATC	743
685			
774	Db	CGTCGTCAACACCCGTCATACGACGAAAGATTAGAAACCATTTAAAAACGTCCAAAAAGC	833
775			
744	Qy	TGGATTTAAAGCGCATGCACAGTGGTATTTCTTGGTCTTGGTGAGACCCCAAGGAGCCGCTGT	803
745			
834	Db	TGGTATCAAAAGCTTGTACCGGTGGTATTTCTCGGATTGGGAAACCCGCAAGAATCATGT	893
835			
804	Qy	ATCTTTCTCTACACCTTTGGCCACAATGGGATCAGCATCCAGAGTCTCTTCCAATCAACAG	863
805			
894	Db	TTCGTTTTTGTACACTTTGTCCAAACATGAGTGCAGCCACCCAGAAATCAATTGCCAATCAACAG	953
895			
864	Qy	ACTGGTCCCAATCAAGGCGACGCCAATGTATGAAGAAGT-----TAAGAACA	911
865			
954	Db	ATTAGTTCCAAATCAGGGTACTCCAAATTTGTGAAGTGTGAAAAATGCCCAAGGAAG	1013
955			
912	Qy	GCAAGTTGAAGTTGATGAGATTGTGAGAACCAATGTCTACTGCAAGATTGGTCAATGCCAAC	971
913			
1014	Db	ACAATTCGAATTTGAAGTCTATTTTGAGACTATTTGCTACTGCTAGATTGATGCTCTGA	1073
1015			
972	Qy	GTCTATTATCAGATTGGCTGCAGGAAGATATACAATGAAAGAGCGACAGAGTGATGTG	1031
973			
1074	Db	ATCCATCATTAGATTGGCCGCGGTGCGATACACCATGAAGGAAACACAGAGCAGTCTTTGTG	1133
1075			
1032	Qy	CTTCATGGCTGGTTGTAATGCCATCTTTCACAGGTGAAGAAATGCTCAACAATGTGTAA	1091
1033			
1134	Db	TTTCATGAGTGGGGTCAATGCTATTTTTCACCGGTGAAGAAATGTTGTACTCAATGTGTAA	1193
1135			
1092	Qy	CGGCTGGGATGAGGATAAAGCCATGTTGGGCTAAATGGGGTCTGAAAACCAATGGAGAGTTT	1151
1093			
1194	Db	TGGTTGGGATGAAGATATCACCATGTTGAAGAAATGGGTTTGAACCAATGGGAAGTTT	1253
1195			
1152	Qy	CA 1153	
1153			
1254	Db	CA 1255	

RESULT 2

US-08-346-849--1
; Sequence 1, Application US/08346849
; Patent No. 5670483
; GENERAL INFORMATION:
; APPLICANT: Zhang, Shuguang
; APPLICANT: Lockshin, Curtis
; APPLICANT: Rich, Alexander
; APPLICANT: Holmes, Todd
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346.849
; FILING DATE:


```

;
; LENGTH: 3083 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1292..2590
; OTHER INFORMATION: /product= "zuotin"
;
US-08-898-300-1

Query Match          27.6%; Score 328.2; DB 4; Length 3083;
Best Local Similarity 72.1%; Pred. No. 2.9e-94;
Matches 445; Conservative 0; Mismatches 163; Indels 9; Gaps 1;

Qy 553 GTCCATCAGCCTTGGAAATGGAGAGTTGTGTCACTCCCTGGGAATGGTTGATATAAGACCAAGCC 612
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 GTGAATGATATGGGGCTAGAAACGTGTGTTACTTTAGTATGGTTGATCAAGATCAAGCA 65

Qy 613 ACTGAAATTGAAAGTGTCTGGGTTGACGGCGTACAAACCAATACATTGATACCTTACAAGAA 672
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 AAGCAATTGAAAGATGCAAGGTTTGACTGTCATACAAACCAATACATCGACACTTCCAGAGAA 125

Qy 673 CACTATCCAAAGGTGATCTCCACAGAGAGCTTTCATGATGATGATGCAAAACATCCAAAAC 732
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 CACTATAGTAAGGTCATCACCACGAGAACCTTACGACGACAGGTTACAGACCATCAAGANT 185

Qy 733 GTTCAAAGGATCTGGATTAAAGGCATGCACAGGTGATATCTTTGGTCTTGGTGAGACCCAA 792
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 GTCCAAGAACTCGAATAAAGCCCTGTACCGTGGTATTTTGGGTCTCGGTGAAGCGAA 245

Qy 793 GAGGACCGTGATCTTTCCTCTCAACCTTGGCGCAATATGATACAGATTCAGAGTCTCTT 852
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 246 GACGACCATATAGGATTCATCTACACATTTATCCAATATGTCTCTCATCTCTGAGTCCCTA 305

Qy 853 CCATCAACAGACTGTTCCCAATCAAGGCGACCAATGTATGAAGAGTT----- 903
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 306 CCAATTAAATAGACTAGTTGCTTATCAAGGGAGTCCATATGGCTGAGGAACCTTGCCGATCCA 365

Qy 904 AAGAAACAAGCAAGTTGAAGTTGATGAGATTGTGCAGAACCAATGCTACTGCAAGATTTGTC 963
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 366 AAGAGTAAAGATTGCAATTCGACGAAATTTTGAGAACCATTTGCCACAGCGAGAATAGTT 425

Qy 964 ATGCCAACGTCATTATCAGATTGGCTGCAGGAAGATATACAATGAAGAGGCGAAGACAG 1023
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 426 ATGCCAAGGCCATTATAAGACTTTGCCGCTGGCTGTATACAATCAAGAAACAGAGACAA 485

Qy 1024 GTGATGTCTTCATGGCTGGTTGTAATGCCATCTTTCAGGTAGAGAAATGCTCACAAACA 1083
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 486 TTTGCTGTCTTCATGGCAGGTTGTAAACAGATATCTTCACCGGTAGAGAAATGCTCAGACA 545

Qy 1084 ATGTGTAACGGCTGGGATGAGGATAAAGCCATGTTTGGCTAATGGGGTCTGAAACCAATG 1143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 546 ATATATAACGGTTGGACGAGAACAGGCAATGTTGGCTAATGGGGATTCACACCTATG 605

Qy 1144 GAGAGTTTCAAAATCAA 1160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 606 GAGGCATTTAAGTACGA 622
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESIST 5

```

US-08-824-513-1
; Sequence 1, Application US/08824513
; Patent No. 6800481
; GENERAL INFORMATION:
; APPLICANT: Holmes, Todd
; APPLICANT: Zhang, Shuguang
; APPLICANT: Rich, Alexander
; APPLICANT: Dipersio, C. Michael
; APPLICANT: Lockshin, Curtis
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; TITLE OF INVENTION: THEREFOR

```

Db 366 AAGAGTAAAGTTGCAATTCGACGAATTTTGAACCAATTTGCCACAGCGAGAAATAGTT 425
Qy 964 ATGCAACAGTCTATTATCAGATTGGCTGCGAGGAGATATACAATGAAGAGGCGAGAACAG 1023
Db 426 ATGCAAGGCAATTATAAGACTTGGCGTGTGTATACATGAAGAAAGAGAGCAA 485
Qy 1024 GTGATGCTTCATGGCTGGTTGTAATGCCATCTTACAGGTAAAGAAATGCTCACAA 1083
Db 486 TTTGTCTGTTTCATGGCGAGTTGTAACAGATATCTTCCCGTAAAGAAATGCTGACGACA 545
Qy 1084 ATGTGTAACGCTGGGATGAGGATAAGCCATGTTGGCTAAATGGGTCTGAAACCAATG 1143
Db 546 ATATATAACGTTGGGAGAGACAGCAATGTTGGCTAAATGGGATTCGAACCTATG 605
Qy 1144 GAGAGTTTCAAAATACAA 1160
Db 606 GAGGCATTTAAGTACGA 622

RESULT 6

US-08-401-068-13
; Sequence 13, Application US/08401068
; Patent No. 5859335
; GENERAL INFORMATION:
; APPLICANT: Patton, David
; TITLE OF INVENTION: Enhanced Biotin Biosynthesis in Plant Tissue
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/401,068
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/351,970
; FILING DATE: 08-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1351 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 59..1192
; OTHER INFORMATION: /product= "Arabidopsis BioB enzyme"

US-08-401-068-13

Query Match 25.6%; Score 304; DB 2; Length 1351;
Best Local Similarity 58.6%; Pred. No. 1e-86;
Matches 551; Conservative 0; Mismatches 380; Indels 9; Gaps 1;
Qy 202 AACACCTGGACCAAGAGAGAAATTAAGCTATATATAGACACACACTCATGGACTTGATG 261
Db 194 AACGATTGGAGTAGAGATGAATCAAGTCTGTTATGATTCCTCTCTTGTGACCTCCTC 253

RESULT 7

US-08-846-338-13
; Sequence 13, Application US/08846338
; Patent No. 5869719
; GENERAL INFORMATION:
; APPLICANT: Patton, David
; TITLE OF INVENTION: Transgenic Plants Having Increased Biotin Content
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5869719artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005

Qy 262 CACTATGCTCAGGTGCAACACACAGAGGTTCCAAACAACCTTTCAGAGGTTCAATTGTGCACT 321
Db 254 TTCATGGAGCTCAGGTTTCATAGACATGTTTCATAAATTCAGGGAGGTACAACAATGTACC 313
Qy 322 CTTATGATATCAAAACTGGTGTGTACCGAGGAGCTAGTACTGTGCGCCATTCACAG 381
Db 314 CTCCTCTCCATAAAGACTGGTGTGTGTAGTGAAGACTGTTTCATATTCTCTCAGTCTTCG 373
Qy 382 CGTTACAACTGGTGTCAAGGCTGAAAGAAATCATCAAGTGTGATGAGGTGATTGAAGCT 441
Db 374 AGATATAGCACTGGAGTTAAGGCACAAAGACTCATGTCTAAGGACGCTGTCATTGATGCT 433
Qy 442 GCAAGAGGCAAGGCAATGGAATCTACAAGTGTCTGTATGGTGTCTGTGGAGAGAG 501
Db 434 GCTAAGAGGCAAGGCAAGGCTGGGAGCACACGTTTTTGTGATGGTGTCTGTGGCGAGAT 493
Qy 502 ATGAAGGTAGAAGTCAAACTTTGAAGAAATCAAGAGATGATCATCTGCTGTCATGAC 561
Db 494 ACAATTGGACGGAACCAACTTCAGCCAGATTTCTTGAATACATCAAGAAATTAAGAGGC 553
Qy 562 CTTTGGATGGAGAGTTGTGTACCCCTGGGAATGTTGATAAAGACCAAGCCACTGAATTG 621
Db 554 ATGGGATGGAGTTTGTCTGCACCTTAGGCATGATTGAGAAACAAAGCACTAGAGCTA 613
Qy 622 AAAAGTCTGGTGTGACGGCTACAACCAATTAACATTGATTAACAAGGAACTATCCA 681
Db 614 AAGAAGCTGGCTCTCACTGCTTATAACCAACAATCTTGTATCTTCAAGAGAGTACTACCA 673
Qy 682 AAGGTATCTCCACAAGAGCTTTGATGATAGATTGAAACATTCAAAACAGTTCAAGGA 741
Db 674 AACGTCTACTACTAGAAAGTTATGACGATCGCCTTGAAACTCTTAGCCATGTTCTGAT 733
Qy 742 TCTGGATTAAAGGCATGCACAGGTGTATTCTTGTGCTTGTGAGACCCCAAGAGGACCGT 801
Db 734 GCTGGAATCAAGTTTGTTCAGGAGGAATCATAGGCTTGTGAGGAGAGGAGACAGA 793
Qy 802 GTATCTTTCTCTACACCTTGGCCACAATGGATCAGCATCCAGAGTCTCTTCCAATCAAC 861
Db 794 ATAGGTTTATTACACACGCTGGCAACACTTCTTCTCACCTCAGAGTGTTCCTCATTAAT 853
Qy 862 AGACTGCTCCCAATCAAGGCGACGCCAATGTATGAAGAGTTAAGAAACAGAGGTTGAA 921
Db 854 GCTTACTTTCAGTGAAGGCACTCTCTTTGAG-----ACCAGAGAGCCAGTTGAG 904
Qy 922 GTTGTAGAGATTGTCAGAACCACTTGTCTACTGCAAGATTGTCATGCAACGCTCTATTATC 981
Db 905 ATATGGAGATGATCAGGATGATTGGAACCGCACGTAATTGTAATGCCAAGGATGGTG 964
Qy 982 AGATTGGCTGCAGGAAGATATACAATGAAGAGGCGAGAAACAGGTGATGCTTTCATGGCT 1041
Db 965 AGACTGCTGCTGGTAGAGTCCGGTTCTCAATGTCCGACACAGCTCTCTGTTTCTTGTCT 1024
Qy 1042 GGTGTAATGCCATCTTTCAGGTAAAGAAATGCTCACAAATGTGTAAACGGCTGGAT 1101
Db 1025 GGTGCAAACTCTATCTTCCCGGAGAGAGCTTTTAAACCACACCAACAATGATTTTGAC 1084
Qy 1102 GAGGATAAGCCATGTTGGCTAAATGGGCTCTGAACCAA 1141
Db 1085 GCTGACCACTCATGTTCAAGACATTAGGCTCATTCCTTA 1124

CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/846,338
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1351 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 59..1192
OTHER INFORMATION: /product= "Arabidopsis biotin synthase enzyme"
US-08-846-338-13

Query Match 25.6%; Score 304; DB 2; Length 1351;
Best Local Similarity 58.8%; Pred. No. 1e-86;
Matches 551; Conservative 0; Mismatches 380; Indels 9; Gaps 1;

QY 202 AACCTGACCAAGAAATTAAGCTATATATGACACCACTCATGACCTTGATG 261
DB 194 AACGATTGAGTAGAGTGAATCAAGTCTGTTATGATCTCTTCTTACCTCTC 253
QY 262 CACTATGCTCAGGTGCAACACAGAAAGTTCCAAACCTTCAAGGTTCAATTTGTCAT 321
DB 254 TTCCATGGAGCTCAGGTTTCATAGACATGTTTCAAACTTCAGGAGGTACAACAATGTACC 313
QY 322 CTTATGAATATCAAACTGGTGTGTCAGGAGCTGTAAGTACTGTGCTCCCAATCACAG 381
DB 314 CTCCTCTCCATAAAGACTGGTGGCTGTAGTGAAGACTGTTTCATATTTGCTCCTCAGTCTTCG 373
QY 382 CGTTACCAACTGGTGTCAAGGCTGAAGAAATCATCAAGTTGATCAGGTGATTGAAGCT 441
DB 374 AGATATAGCACTGGAGTTAAGGCAAAAGACTCATGTCCTAAGGAGCTGTCAATGATGT 433
QY 442 GCAAGAGGCAAGGCAATGATGCTCAAGGTTCTGTATGGTGTCTGCTTGGAGAGAG 501
DB 434 GCTAAGAGGCAAGGCAAGGCTGGAGCAACGTTTGTGATGGTGTCTGCTTGGAGAGAT 493
QY 502 ATGAAAGGTAGAAAGTCAAACTTGAGAAATCAAAAGAGATGATCACTGCTGCCATGAC 561
DB 494 ACAATGGAGGCAAGGCAAGGCTGAGGCAAGTCTGAGGAGTCTGAGGAGTCTGAGGAGC 553
QY 562 CTTGGAATGGAGATTGTGTCACCTTGGGAATGTTGATAAAGCAAGCAAGCACTGAATTG 621
DB 554 ATGGGAGTGAAGTTGCTGCACTTGGAGTATGATGAGGAGTGAAGGAGTGAAGGAGT 613
QY 622 AAAAGTGTGGTGTGACGGGTAACCAATGATGATGATGATGATGATGATGATGATGATGAT 681
DB 614 AAGAAGGCTGGCTCACTGCTTATAACCAAACTTGTGATGATGATGATGATGATGATGATGAT 673
QY 682 AAGTGTATCCACAGAGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 741
DB 674 AACGTCATCACTACTAGAAAGTTATGACGATCGCTTGAAGAACTCTTAGCCATGTTGCTGAT 733

QY 742 TCTGGATTAAAGGCATGCACAGGTGGTATTCTTGGTCTTGTGTGAGACCAAGAGGACCT 801
DB 734 GCTGGAATCAACGTTTGTTCAGGAGGAATCATAGGCTTGGTGTGAGGAGGAGGAGACAGA 793
QY 802 GTATCTTTCTCTACACCTTGGCCCAATGGATCAGCATCCAGAGTCTCTTCCCAATCAAC 861
DB 794 ATAGGTTTATTACACACGCTGGCAACACTTCTTCTCACCTTGAGAGTGTCCCATTAAT 853
QY 862 AGACTGGTCCCAATCAAGGCGACGCAATGTATGAAGAAGTTAAGAAACAAGCAAGTTGAA 921
DB 854 GCTCTACTTGTGAGTGAAGGCACTCTCTTGAAG-----ACCAGAGGCAAGTTGAG 904
QY 922 GTTGATGAGATTGTGAGAACCAATCTCTACTGCAAGATTGGTCTCATGCCAACGTTCTATTATC 981
DB 905 ATATGGGAGATGATCAGGATGATTGGAACCGCACCTGATTGTAATGTCAAAAGCGATGGTG 964
QY 982 AGATTGGCTGCAGGAAGATATACATGAAGAGGCGAGACAGGTGATGTGCTTCATGGCT 1041
DB 965 AGACTGTCTGCTGTAGAGTCCGTTCTCAATGTCCGAACAAGCTCTCTGTTTCTTCTTGT 1024
QY 1042 GGTTCGTAATGTCATCTTCAAGGTAAGAAATGCTCACAACAATGTGTAACCGCTGGGAT 1101
DB 1025 GGTGCAAACTCTATCTTCCCGGAGAGAGCTTTTAAACCAACCAACAATGATTTGAC 1084
QY 1102 GAGGATGAAGCAATTTGGCTAAATGGGCTCGAAACCAA 1141
DB 1085 GCTGACCAGCTCATGTTCAAGACATTAGGCTTCATTCCTA 1124

RESULT 8
US-09-328-352-2698
Sequence 2698, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 2698
LENGTH: 1017
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-2698

Query Match 19.9%; Score 236; DB 4; Length 1017;
Best Local Similarity 54.7%; Pred. No. 5.9e-65;
Matches 520; Conservative 0; Mismatches 415; Indels 15; Gaps 2;

QY 130 AAGACACAGCTCAACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTC 249
DB 28 ATGACTCTAGCTAATGATGATGAGCTCGTGAAGAAATCCAAAGCTCTATATGAACAACCTTTT 87
QY 250 ATGGACTTGTGACACTATGCTCAGGTGCAACACAGAGGTTTCCAAACACCTTCAGAGGTT 309
DB 88 TTGGATTTAGTTTTTAAGGCTCAACAGTACACCGTGAGCACTTCACTGCCAATACAATT 147
QY 310 CAATTGTGACCTCTTATGAATATCAAACTGTGTGTGTTGTACCGAGAGCTGTAAGTACTGT 369
DB 148 CAGGTGACGACCCCTTTTATCTATTAAAACGGGTAAATGCTCTGGAAGATTGCAAAATATTGC 207
QY 370 GCCCAATCACAGGTTTCAACACCTGTGTCAGGCTGAAAGAAATCATCCAAGTTTCATGAG 429
DB 208 TCTCAATCTGCACATTACGATTCAAAATTAGAAGCAGAAATAACGTTATTGCTGTTGAANA 267
QY 430 GTGATTGAAGTGCAGGAGGCAAGGCAATGATGATGATGATGATGATGATGATGATGATGAT 489
DB 268 GTAATTAGCGAAGCAAAAGCTGCAAAAGATTCAAGTTTCATCTCGTTTTTGTATGGTGTCT 327
QY 490 GCTTGGAGAGATGAAGGTTGAAGGTTGAAGGTTGAAGGTTGAAGGTTGAAGGTTGAAGGTT 549

Db 328 CTTGGCG-----CAACCCCTCATGAGCGGATATGCTTATGTTTTAGAAATGGTACGT 381
Qy 550 GCTGTCATGACCTTGAATGAGAGTGTGTCAACCTCGGAATGGTGTGATAAGACCAA 609
Db 382 GAAGTTAAGCATTAGGTATGGAACCTGTATGACTTTAGTATGCTTATCATCAACAA 441
Qy 610 GCCACTGAATGAAAGTGTGGTTGACGGCGTACCAACCAATATGATGATCAATCAAG 669
Db 442 GCAGAACGCCCTAAAGATGCGGTTAGACTATTACAAACCAATACCTTGATCTCTCGC 501
Qy 670 GAACACTATCAAGGTGATCTCCACAAGAGCTTTGATGATAGATGAAACATTCAAA 729
Db 502 GAGTATTAATTTCCCAATCATCAGACACAGTACTTTTATGATGACCGTTTGAATACTTAGAT 561
Qy 730 AACCTTCAAGGATCTGGATTAAGGCAATGACAGGTGTATCTTGTGCTTGGTGAGACC 789
Db 562 TATGTAGCTCAAGCTGGTATGAAGTATGATGCGCGTATTTGTTTATGTTAGGTGAGGC 621
Qy 790 CAAGAGGACCGTGTATCTTCTCTACACCTTGGCCCAATGATGATCAGCATCCAGAGTCT 849
Db 622 CGTGAAGATCGCATTTGCTTATACATGAGTTAGCAACCTTACCTATTATCCAGATCT 681
Qy 850 CTTCCAAATCAAGACTGTGTCCTCAATCAAGGCGACGCCAATGATGATGAAGAGTTAAGAAC 909
Db 682 GTGCCAATCAACATGCTTGTTCCAATTAAGGTACACCAATAGCTAGTGTG----- 733
Qy 910 AAGCAAGTTCAAGTTGATGATGATGTCAGAACCAATGCTACTGCAAGATTGGTCATGCCA 969
Db 734 -AAACTAGATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 792
Qy 970 ACCTGTTATTCAGATTTGCTGAGGAGATATACAAATGAAGAGGACGACAGGTGATG 1029
Db 793 CATAGTATATTCGTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852
Qy 1030 TGCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1089
Db 853 GCATTTATGCGGTGCAAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
Qy 1090 AACGGCTGGATGAGGATAAGCCATGCTGGCTAAATGGGCTCTGAAACC 1139
Db 913 GCAGGTGAAGTTAAGACCAAGCGTTATTTAATAAATTAGGTTAAGC 962

RESULT 9

US-09-902-540-7391
; Sequence 7391, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 7391
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-7391

Query Match 19.4%; Score 230.4; DB 4; Length 954;
Best Local Similarity 54.5%; Pred. No. 3.6e-63;
Matches 512; Conservative 0; Mismatches 416; Indels 12; Gaps 2;
Qy 207 CTGGACCAAGAGAATTAAGCTATATATGACACCACTCATGACTTATGACTTA 266
Db 9 CTGGTCGCTGCGAGGTAGGCGCCCTCTACGAACCTCCCGCTGCTGACCTGCTGACAA 68

Qy 267 TGCTCAGGTCAACACACAGAAAGGTTCCAAACAACTTCAGAGGTTCAATTTGTCATCTTAT 326
Db 69 GGCCACAGACGGTCCACCGGCCCTGTTGCTGGGCAACACAGGTCAGCTCTGTTGCTGCT 128
Qy 327 GAATATCAAAACTGGTGTGTTACCGAGGAGTGTAACTGTTGCTGCTGCTGCTGCTGCTGCT 386
Db 129 GTCCATCAAGACGGGCGGCTGCCCCGAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 188
Qy 387 CAACACTGGTGTCAAGGCTGAAAGAAATCATCAAGTTGATGAGGTGATTTGAAGCTGCAAA 446
Db 189 CAAGACGGGCTCAAGGCGGAGAGCTGATGGCGTCCCGAGCTGCTGCTGCTGCTGCTGCT 248
Qy 447 GAGGCAAGGCGCAATGGATCTTCAAGGTTCTGTATGGTGTGCTGCTGCTGCTGCTGCTGCT 506
Db 249 CAAGGCGCGCGCGCGAGCCACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 308
Qy 507 AGTAGAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGCTGCTGCTGCTGCTGCTGCT 566
Db 309 GAGCGG---TCCGCGAGTTCCAGACGCTGCTGAGATGCTGCGCGCTGCGCGCTGGG 365
Qy 567 AATGGAGAGTTGTGTCACCTCGGAATGTTGATAAGACCAAGCCACTGAATTTGAAAG 626
Db 366 CATGGAGCGGTGTCACCTCGGCGATGCTTCCGAGAGCAGCGGAGCGCTTCCGCGA 425
Qy 627 TGCTGGGTTGACGGCGTACCAACATTAACATTTGATATCTTCAAGGAAACACTATCCAAAGGT 686
Db 426 GCGCGGCTGTCGCGGTACCAACCAACCTGGACAGCTGCGCGGAGCACTACGCGGACAT 485
Qy 687 GATCTCACAAAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 746
Db 486 CATCTCCACCGCACCTTATGAAGACCGCTGCGCACGCTCAACCGCTGCGCGGCGCGG 545
Qy 747 ATTAAAGGCATGACAGGTGTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806
Db 546 CATCTCGGTGCTGCGCGGCGATCAATGSCATGGGCGAGTGGTGTGATGACCGCTGCA 605
Qy 807 TTTCTCTACACCTTGGCCCAATGGATCAGCATCCAGAGTCTCTTCCAAATCAACAGACT 866
Db 606 CTTGCTGCGCACCTTGGCCCAACAGGAGCACACCGGAGTGGTGGCTTCCCATCAACGGCT 665
Qy 867 GGTCCCAATCAAGGCGACCGCAATGATGAGAGATTTAAGAACAGCAAGTTGAGTTGA 926
Db 666 GTGGGCGGTGGAGGCGACCGCTGCGAGG-----AGCAGCAGCGCGTGGAGACGGT 716
Qy 927 TGAGATTGTCAAGAACCATTTGCTACTCAAGATTGCTCATGCCAACGCTCTATTATCAGATT 986
Db 717 GGACATGGTGGCACCATTCGCCCGCGCATCTCTGATGCCCCAGTCTGTCATGATGGCGCCT 776
Qy 987 GGCTGAGGAAGATATACAAATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1046
Db 777 GTCCGCGGCGCGGACGAGATGAACGAGAGGCGCAACTGCTGTCATGATGGCGGCGC 836
Qy 1047 TAATGCACTTTCAAGGTAAGAAATGCTCAACAAATGTTGTAACGGCTGGGATGAGGA 1106
Db 837 CAACTCACTCTTCTTCCGCGAGAGCTGCTACCAACCGGCAACCCCGAGTACACCCAGGA 896
Qy 1107 TAAGGCATGTTGGCTTAATGGGCTGTAACCAATGGAG 1146
Db 897 CATGGCCCTGCTGGAGAGGCGGGAATCCGCCCCCTGGAG 936

RESULT 10

US-09-902-540-709/c
; Sequence 709, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 709
; LENGTH: 4183
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-709

Query Match 19.4%; Score 230.4; DB 4; Length 4183;
Best Local Similarity 54.5%; Pred. No. 9.1e-63;
Matches 512; Conservative 0; Mismatches 416; Indels 12; Gaps 2;
QY 207 CTGACCAAGAGAAATTAAAGCTATATATGACACACCACTCATGACCTTGTGCACTA 266
DB 2127 CTGGTCGCTGCGAGGTGAGGCGCTCTACGAACCTCCGCTGCTGGACCTGTGACAA 2068
QY 267 TGCTCAGGTGCAACACAGAGGTTCCAAACACCTTCAGAGGTTCAATTTGCACTCTTAT 326
DB 2067 GGCCAGACGGTCCACCGGCGCTGTTCTGGACACAAAGTCCAGCTCTGTTGCTGCT 2008
QY 327 GAATATCAAAACTGGTGGTGTACCGAGGACTGTAAGTACTGTGCCCAATCAAGCGTTA 386
DB 2007 GTCCATCAAGACGCGCGCTGCCGAGGAGTCTCTGTAAGTCCGCGAGCGCTGCTA 1948
QY 387 CAACACTGGTCAAGGCTCAAGAAATCATCAAGTTGATGAGTGTGATTTGAAGCTCAAA 446
DB 1947 CAAGACGGGCGTCAAGCGGAGAGCTGATGGCGGTGCCGAGCTGCTGATCCGCGTC 1888
QY 447 GGAGGCAAGGCAATGATCTCAAGGTTCTGTATGGGTGCTGCTTGGAGAGAGATGAA 506
DB 1887 CAAGCGCGCGCGCGGAGACCCGCTTCTGATGGGCGCGCTGGCGGAGGTGAA 1828
QY 507 AGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGTCCATGACCTTG 566
DB 1827 GGACGG---TCCGAGTTCCACAGCGTGTGGAGATGTTGCGGCGCTGCGCGCTGG 1771
QY 567 AATGGAGAGTTGTCTACCGCTGGATGTTGATTAAGACCAAGCCACTGAATTTGAAG 626
DB 1770 CATGGAGGCGTGTGCCACCTTGGCATGCTCTCCGAGAGCCAGGCGCTGCGCGA 1711
QY 627 TGCTGGTTGACGCGGTACCAACCATTAACATTTGATCTTCAAGGAACTATCAAAAGT 686
DB 1710 GGCGGCGCTGTCCGCTACCAACCACTGGACAGTCCCGCGAGCACTACGCGCAT 1651
QY 687 GATCTCCAAAGAGCTTTGATGATGATTAAGAAATCAAAAGCTTCAAGGATCTGG 746
DB 1650 CATCTCCACCGCACTTATGAAGACCGCTTCCGACGCTCAACCGCGTGGCGACGCG 1591
QY 747 ATTAAGGCATGCACAGGTGTTCTTGGTCTTGGTGAGACCCAGAGACCGTGTATC 806
DB 1590 CATCTCGTGTGTGGCGGCGCATATGGCATGGGCGAGTCCGGTGGATGACCGTGCAA 1531
QY 807 TTTCCCTTACACCTTGGCCCAATGATGATGATGATGATGATGATGATGATGATGAT 866
DB 1530 CTGCTCGGACCTTGGCCCAACAGGAGACACCGCGAGTCCGTGCGCATCAACGCT 1471
QY 867 GTTCCCAATCAAGGCGACGCAATGATGATGATGATGATGATGATGATGATGATGAT 926
DB 1470 GGTGGCGGTGGAGGCGACGCGCTGACG-----AGCAGACGCGCTGGAGACGCT 1420
QY 927 TGAGATGTCAGAACCATTTGCTACTGCAAGATTGGTTCATGCCAACCTCTATTATCAGAT 986
DB 1419 GGACATGGTGGCAGCATTCATCCGCGCGCATCTGATGCCCGAGTCAATGTCGCT 1360
QY 987 GGCTGAGGAGATATACATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1046
DB 1359 GTCCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1300
QY 1047 TAATGCCATCTTCAAGGTAAGAAATGCTTCAACCAATGTTAAAGGCTGGGATGAGGA 1106
DB 1299 CAACCTCACTTCTTGGCGGAGAGCTGCTCAACCGGCAACCCCGAGTACACCCAGGA 1240

QY 1107 TAAAGCCATGTTGGCTAAATGGGCTCTGAAACCAATGGAG 1146
DB 1239 CATGGCCCTGCTGGAGAGCGGGAATCCGCCCTGGAG 1200

RESULT 11

US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581

GENERAL INFORMATION:

APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB186P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 17.8%; Score 211.6; DB 4; Length 1830121;

Best Local Similarity 54.3%; Pred. No. 4.7e-55;

Matches 500; Conservative 0; Mismatches 404; Indels 16; Gaps 3;

QY 221 AAATTAAGCTATATATGACACACCACTCATGCACTTGTGCACTATGCTCAGGTCAAC 280
DB 1085608 AAGTTGAGCGGTATTGAAAGCCATTTTGGATTAGTTTACCGAGCAACGCAAGTTC 1085667
QY 281 ACAGAGGTTCCAAACACCTTCAGAGGTTCAATTTGCACTCTTTATGAATATCAAACTG 340
DB 1085668 ATCGTAAGCATTTTAATCTCTCGCGGATTCAGTTATCCACGTTAATGTCTATCAAAACGG 1085727
QY 341 GTGGTTGTACCGAGCACTGTAGTACTGTGCCAATCACAGGTTTACACACTGGTGCA 400
DB 1085728 GGGGATGCCAGAGATTGTAGTTATTGCCCTCAATCAGCCCGTTATCATCTGCGGTAC 1085787
QY 401 AGCGTGAAGAATCATCCAAAGTTGATGAGTGAAGCTGCAAGCTGCAAGAGGCAAGGCCA 460
DB 1085788 AAATCAGCAGTTATTAGATGTTGATGAGATCGTCGCTAAAGCAAAATTTGCAAGAC 1085847
QY 461 ATGGATCTCAAGGTTCTGTATGGTGTGCTGTTGGAGAGAGATGAAGGTGAAGAGTCAA 520

```
Db 1085848 GTGTGTCAGGGGTTTCTGTATGGGGCTGTCTGGCGAGGCCCTAAGCCAA-----AAG 1085901
Qy 521 ACTTGAAGAAATCAAGAGATGATCACTGTGTCTCATGACCTTGGAAATGAGAGTTGTG 580
Db 1085902 ATATTGAGAAAGTCACAGAAATTAATAAGCCGTGAATCACTTGTGTTAGAACTTCGG 1085961
Qy 581 TCACCTCGGAAATGGTTGATGAAGACCAAGCCACTGAATGAAAGTCTGGTTGACGG 640
Db 1085962 GTACCTTTGGTTTATTGCAAGACGGTATGGCGGAAGATTAAAGAAAGCTGGATTAGATT 1086021
Qy 641 CGTACACCAATAACATTGATATTACAGGACACTATCCAAAGGTGATCCCAAGAA 700
Db 1086022 ATTACAAATCAATCTTGATAC-CGCACAGAAATTAAGCAAGTGAATGGTACTCGCC 1086080
Qy 701 GCTTTGATGATAGATTGAAACATCAAAAGCTTCAAGGATCTGGATTAAGGATGCA 760
Db 1086081 GTTTTGATGATCGCTTAGCACCTTAGGAAAGTGGCAAGCTGGATTAAGAGTGTG 1086140
Qy 761 CAGTGGTGTATCTTTGGTCTTGGTGAGACCCCAAGAGCCGTGTATCTTTCTCTACACT 820
Db 1086141 GTGTGGCATTTGTAGGGATGAATGAACCTCGCAAGAACGGCTGGATTAAATCGGAGCC 1086200
Qy 821 TGGCCACAAATGATCAGATCAGAGTCTCTTCCAAATCAACAGACTGTGCCAATCAAGG 880
Db 1086201 TTGCAAAATCTTGATCCGCAACCTGAGTCAGTCCGCAATTAATCAACTGTTAAGGTTGAAG 1086260
Qy 881 GCAGCCCAATGATGAAGAGTTAAGAACAGCAAGCTTGAAGTTGATCAGATTGTGCAAA 940
Db 1086261 GTACGCCCTTAGCCGATCGGG-----AGAAATTAGATTGACGAGGTTTGTGCGAA 1086311
Qy 941 CCATTGTCTACTGCAAGATTGGTCTATGCGCAAGCTCTATTATCAGATTGGCTGCAGGAAGAT 1000
Db 1086312 CGATTGTCTTGGCAGTATTACTATGCCCCAAAAGTTATGTTGCGCTTCTGCAGGTGCA 1086371
Qy 1001 ATACAATGAAGAGCAGCAAGAGTGATGCTTCAATGCGCTGGTGTGTAATGCCATCTTCA 1060
Db 1086372 GTGGTATGACTGAAGAAATGCAAGCGATGTGTTTATGCGGGTGCAGAAATCTATTTTCT 1086431
Qy 1061 CAGGTGAAGAAATGCTCACAACAATGTGTAACGGCTGGGATGAGGATAAAGCCATGTTGG 1120
Db 1086432 ATGGCGATAATTACTTGTACGATAATCCAGAAAGATGGCGATCATGTTATGATGG 1086491
Qy 1121 CTAATGGGGTCTGAAACCA 1140
Db 1086492 CAAAATTAGATTAGAGCCA 1086511
```

RESULT 12

US-09-643-990A-1

Sequence 1, Application US/09643990A

Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann

Mark D. Adams

Owen White

Hamilton O. Smith

J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville,

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/643.990A

FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787

FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PBI86P1C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-610-5790

TELEFAX: 310-309-8439

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1

Query Match

17.8%; Score 211.6; DB 4; Length 1830121;

Best Local Similarity 54.3%; Pred. No. 4.7e-55;

Matches 500; Conservative 0; Mismatches 404; Indels 16; Gaps 3;

Qy 221 AAATAAGCTATATATGACACACCTCATGGACTATGTCAGTCTGTCAGTGCAC 280

Db 1085608 AAGTTGAGGGCTTATTTGAAACGCCATTTTGGAAATTTTACCGAGCAACGCAAGTTC 1085667

Qy 281 ACAGAAGTTCCAAACCTTCAGAGTTCAATTTGTCACCTTTATGAATATCAAACTG 340

Db 1085668 ATCTAAGCATTTTAATCTTCGCGCGATTCACTTATCCAGTTATGTCATCAAAACGG 1085727

Qy 341 GTGCTTGTACCGAGACTGTAAGTACTGTGCCCAATCAGAGCTTACAACTGCTGTGTC 400

Db 1085728 GGGGATCCCAAGAGATTGATTGTTTATGCTCAATCAGCCCGTTATCATCTGCGGTAC 1085787

Qy 401 AGGTGAAGAAATCATCCAGTTGATGAGGTGATGAAGCTGCAAGAGGCAAGGCCA 460

Db 1085788 AAAATCAGCAGTTATTAGATGTTGATGAGATCGTCGCTAAAGCAAAAATTCGAAAGCAC 1085847

Qy 461 ATGGATCTACAAGGTTCTGTATGGTCTGCTTGGAGAGATCAAAAGGTAGAAAGTCAA 520

Db 1085848 GTGGTGAGGGCGTTTCTGTATGGGGCTGCTTGGCGAGGCCCTAAGCCAA-----AAG 1085901

Qy 521 ACTTGAAGAAATCAAGAGATGATCACTGTGTCTCATGACCTTGGAAATGAGAGTTGTG 580

Db 1085902 ATATTGAGAAAGTCACAGAAATTAATAAGCCGTGAATCACTTGTGTTAGAACTTCGG 1085961

Qy 581 TCACCTCGGAAATGGTTGATGAAGACCAAGCCACTGAATGAAAGTCTGGTTGACGG 640

Db 1085962 GTACCTTTGGTTTATTGCAAGACGGTATGGCGGAAGATTAAAGAAAGCTGGATTAGATT 1086021

Qy 641 CGTACACCAATAACATTGATGATCTTACAGGAACTATCCAAAGGTGATCTCCCAAGAA 700

Db 1086022 ATTACAAATCAATCTTGATAC-CGCACAGAAATTAAGCAAGTGAATGGTACTCGCC 1086080

Qy 701 GCTTTGATGATAGATTGAAACATTTCAAAACGTTCAAGGATCTGGATTAAGGATGCA 760

Db 1086081 GTTTTGATGATCGCTTAGCACCTTAGGAAAGTGGCAAGCTGGATTAAGAGTGTG 1086140

Qy 761 CAGTGGTGTATCTTTGGTCTTGGTGAGACCCCAAGAGCCGTGTATCTTTCTCTACACT 820

Db 1086141 GTGGTGGCATTTGAGGATGAATGAACCTCGCAAGAACGGGCTGGATTAAATCGGAGCC 1086200

Qy 821 TGGCCCAATGATCAGATCAGAGTCTCTTCCAAATCAACAGACTGTGCCAATCAAGG 880

Db 1086201 TTGCAAAATCTTGATCCGCAACCTGAGTCAGTCCGCAATTAATCAACTGTTAAGGTTGAAG 1086260

```
Qy 881 GCACGCCAATGTATCAAGAAGTTAAGACAAGCAAGTTGAAGTTGATGAGATTGTCAGAA 940
Db 1086311 GTACGCCGTAGCCGATGCGG-----AAGATTAGATTGGACGAGTTTGTGCGAA 1086311
Qy 941 CCATTGCTACTCAAGATGTCATGCAACGCTCTATTATCAGATTGGCTGCAGCAAGAT 1000
Db 1086312 CGATTGCTGTTGACGATTAATCTACTATGCCCCAAAGATTATGTCGCTTTCTGCAGGTCGCA 1086371
Qy 1001 ATACAATCAAGAGCGCAGACAGGTGATGTCCTTCATGCTGGTGTGTAATGCCATCTTCA 1060
Db 1086372 GTGTATGACGTAAGAATAAGCAAGGATGTTTATGCGCGGTGCATAATCTATTCTT 1086431
Qy 1061 CAGGTAAAGAAATGCTCACAAATGTGTAAACGCTGGGATGAGGATAAAGCCATGTGG 1120
Db 1086432 ATGCGGATAAATTACTTGTTCAGATTAATCCAGAAGAGATGGCGATCAGTTATTGATGG 1086491
Qy 1121 CTAATGGGCTGTAACCA 1140
Db 1086492 CAAAATTAGATTAGAGCCA 1086511

RESULT 13
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (30001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (330001)..(345000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (345001)..(360000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (360001)..(375000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (375001)..(390000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (390001)..(405000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (405001)..(420000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (420001)..(435000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (435001)..(450000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (450001)..(465000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (465001)..(480000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (480001)..(495000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (495001)..(510000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (510001)..(525000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (525001)..(540000)
```

```
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (540001)..(555000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (555001)..(570000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (570001)..(585000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (585001)..(600000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (600001)..(615000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (615001)..(630000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (630001)..(645000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (645001)..(660000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (660001)..(675000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (675001)..(690000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (690001)..(705000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (705001)..(720000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (720001)..(735000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (735001)..(750000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (750001)..(765000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (765001)..(780000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (780001)..(795000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (795001)..(810000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (810001)..(825000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (825001)..(840000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (840001)..(855000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (855001)..(870000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (870001)..(885000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (885001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
```

```
; NAME/KEY: misc feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature

Query Match 17.5%; Score 208.2; DB 4; Length 1230025;
Best Local Similarity 53.5%; Pred. No. 4.5e-54;
Matches 488; Conservative 0; Mismatches 413; Indels 12; Gaps 2;

Qy 206 CCTGGACCAAGAAGAAATTAAGCTATATATGACACACCACTCATGAGCTTGATGCACT 265
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208824 CCTGGTCATTAGAAGACATCGCGAAATTTATACACTCCGTTATTTGAGCTGATTCACA 1208765
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 266 ATGCTCAGGTGCAACACAGAGGTTCCAAACACTTCAGAGGTTCAATTGTCGACTCTTA 325
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208764 AAGCCAATGCCATATTGCGTAGTAATTTCTCCATTGAGAACTGCAGACTTGCCTATCTGA 1208705
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 326 TGAATATCAAAACTGGTGGTTGTACCCAGGACTGTAACTGATGAGGTTGATGAGCTGCA 385
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208704 TTTGATTTAAACTGGTGGATGCGTTGAAGATTGCGCTACTGTGCCCAATCTTTCCCGCT 1208645
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 386 ACAACACTGGTGTCAAGGCTGAAAGAATCATCCAAAGTTGATGAGGTTGATGAGCTGCA 445
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208644 ATCATACCAGTCACACAGAACCTATGATGAATTTAGAGCTTTGTGGAAGGGCAA 1208585
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 446 AGGAGCAAGGGCAATGGATCTACAAGGTTCTGTATGGGTGCTGCTTGGAGAGAGATGA 505
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208584 AACGTGCTGTAGAGTAGGCGCCACTCGTGTGTCTTTGGGGCTGCTGGCGCAATGCTA 1208525
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 506 AAGGTAGAAAGTCNAACTTGAAGAAATCAAGAGATGATCACTGCTGTCATCACTGCTTG 565
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208524 AGG---ACGATCGATACCTTTGATAGAGTCTCGCTATGTTGTTGAAAGTATCACAGCTCG 1208468
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 566 GAATGGAGAGTTGTGTCACCTCGGAATGTTGATAAGACCAGCCACTGAAATTGAAA 625
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208467 GAGCCGAGGTTGTTGTGCTTTAGGCATGCTCTCCGAGAGCAGCTTAAAAAAGCTGTATG 1208408
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 626 GTGCTGGTTGACGGGTACAACCATTAACAATTGATCTTCAAGGAACACTATCCAAAGG 685
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208407 ATGCAGGACTTTATGCTCAATCAATAATTTAGACTCTTCTCGGAATTTCTATGAAACTA 1208348
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 686 TGATCTCCACAAGAGCTTTGATGATGATGATGAAACATTCAAAACGTTCAAGAGTCTG 745
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208347 TAATCAACAACAGCTTCTTATGAAGATGCGCTCAACACTCTTGATGTAGTAAATAAATCTG 1208288
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 746 GATTAAAGGCATGCACAGGTGGTATTCTTGGTCTTGGTGAGACCCAGAGAGACCGTCTAT 805
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208287 GCATTAGTACATGCTGGGTGGTATTGTAGGTATGGAGAAATCTGAAGAAGACCGTATAA 1208228
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 806 CTTTCTCTACACCTTTGGCCACAATGGATCAGCATCCAGAGTCTCTTTCCAATCAACAGAC 865
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208227 AGCTTCTTCATGTTCTTGGCAACAAGAGATCATATCCAGAATCCGTACTGTAAATTTAC 1208168
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 866 TGGTCCCAATCAAGGGCAGCCCAATGTATGAAGAAGTTAAGAACAAGCAAGGTTGAAGTTG 925
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208167 TTTGGCGGATTTGACGCGACGCTTTTGCAGACACCGCTCCGA-----TTTCTTTCT 1208117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 926 ATGAGATTGTCAACAACATTTGCTACTGCAAGATTGGTCATGCCCAACGCTCTATTATCAGAT 985
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208116 GGGAGTCTTCCGAACCATAGCAACGCGACGGTGTGTTTCCCAAGATCCATGTGACGAC 1208057
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 986 TGGCTGCAGGAAGATATACAATGAAAGAGGACAGACAGGTGTGTGCTTCTCATGGCTGGTT 1045
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1208056 TTGCTGCAGGACGCGCTTTCTCCACAGTAGAACAACAACCTTATGTTTTTCTAGCCGGTG 1207997
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1046 GTHATGCCATCTTCACAGGTAAAGAAATGCTCAACAATGTGTAAAGCGGTGGGATGAGG 1105
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1207996 CCAACTCCATATTCTATGGAGATAAACTGTTGACTGTAGAAAAACAATGATATAGATGAAG 1207937
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1106 ATAAAGCCATGTT 1118
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1207936 ATGCTGAAATGAT 1207924
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 14

US-09-438-185A-1/c
; Sequence 1, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1230230
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-438-185A-1

Query Match 17.5%; Score 208.2; DB 4; Length 1230230;
Best Local Similarity 53.5%; Pred. No. 4.5e-54;
Matches 488; Conservative 0; Mismatches 413; Indels 12; Gaps 2;

QY 206 CTTGGACCAAGAAAGTAAAGCTATATATGACACACACCTCATGACCTTGTGACCT 265
DB 1198672 CTTGGTCAITGAAGACATCCGGAATTTATCACCTCCGTTATTGAGCTGATTCACA 1198613

QY 266 ATGCTCAGGTGCAACAGAGGTTTCAACAACCTTCAGAGGTTCAATTTGTGACCTTTA 325
DB 1198612 AAGCCAAATGCTATTTGCGTAGTAATTTCTCCATTGAGAACTGCAGACTTGTCTGTA 1198553

QY 326 TGAATATCAAACTGTGTGTTTACCGAGACTGTAAGTACTGTGCCAATCACAGGTT 385
DB 1198552 TTTTCGATTAACCTGTGGTAGTGGTTGAAGATTGCGCTACTGTGCCAATCTTCCCGCT 1198493

QY 386 ACAACACTGTGTCAAGCTGAAAGATCATCAAGTTGATGAGGTGATTGAAGCTGCAA 445
DB 1198492 ATCATCCACGTCACACAGAACCTATGATGAATTTGACGTTGTGGAAGGCGCA 1198433

QY 446 AGGAGGCAAGGCCAATGGATCTACAAAGGTTCTGTATGGGTGCTGTGGAGAGATGA 505
DB 1198432 AACGTGCTGTAGAGCTAGGCGCCACTCGTGTGTGTCTTGGGGCTGCTGGCGCAATGCTA 1198373

QY 506 AAGGTAGAAGTCAAACTTGAAGAAATCAAGAGATGATCATCTGCTTCATGACCTTG 565
DB 1198372 AGG---ACGATCGATATCTTTGATAGAGTCTCCTGATGTTGTAAGATATCACAGATCTCG 1198316

QY 566 GAATGAGAGTCTGTCCACCTCGGAATGTTGATAAAGCAACCAAGCACTGAAATTCGAAA 625
DB 1198315 GAGCCAGGTTTGTGTCTTTAGGACATGCTCTCCGAGAGCAAGCTAAAACTGTATG 1198256

QY 626 GTGCTGGTTGACGGGTACAAACATAAATGATATCTTACAAAGGAACACTATCCAAAGG 685
DB 1198255 ATGCAGGACTTTATGCTCAATCAATCAATTAATTTAGACTCTTCTCCGGAATTTCTATGAAACTA 1198196

QY 686 TGATCTCCCAAGAGCTTTGATGATAGATTGAACATTTCAAAACGTTCAAGGATCTG 745
DB 1198195 TAATCAACACAGCTTTCTATGAAGATGCGCTCAACACTCTTGTATGATGATTAATAATCTG 1198136

QY 746 GATTAAGGATGACAGAGGTGTTATCTTGGTCTTGTGAGAGCCCAAGAGGACCGTGTAT 805
DB 1198135 GCATTAGTACATGCTCGGTGTTTGTAGTATGGAGNATCTGAAGAGACCGTATTA 1198076

QY 806 CTTTCCTCTACACCTTGGCCACAAATGGATCAGCATCCAGAGTCTCTTCCAATCAACAGAC 865

DB 1198075 AGCTTCTTCATGTTCTTTCACCAAGAGATCATATCCAGAATCCGTACCTGTAAATTTAC 1198016
QY 866 TGTGCCCAATCAAGGCGCAGCAATGTATGAAGAGTTAAGAACAGCAAGTTGAAGTTG 925
DB 1198015 TTTGGCCGATTGACGCGCAGCTTTTGAAGACCAAGCTCCGA-----TTTCTTTCT 1197965
QY 926 ATGAGATTCTCAGAACCATTGCTTACTTGAAGATTGGTCATGCGCAACGTTCTATTATCAGAT 985
DB 1197964 GGAAGTCTTGGCAACCATAGCAACGGCAGCGGTGTTTTTCCCGAGATCCATGGTACGAC 1197905
QY 986 TGGCTCAGAGAGATATACAAATGAAGAGGAGAGAGAGTGTCTTCTTATGCGCTGGTT 1045
DB 1197904 TTGCTGACGAGCGCGCTTTCTCAGAGTAGAACAACAACCTTATGTTTCTTAGCCGGTG 1197845
QY 1046 GTAATGCCATCTTACAGGTAAAGAAATGCTCACAACAATGTGTAACGGCTGGGATGAGG 1105
DB 1197844 CCACTCCATATCTATGGAGATAAACTGTTGACTGTAGAAAACATGATATAGATGAAG 1197785
QY 1106 ATAAAGCCATGTT 1118
DB 1197784 ATGCTGAAATGAT 1197772

RESULT 15

US-09-180-109A-23
; Sequence 23, Application US/09180109A
; Patent No. 6410293
; GENERAL INFORMATION:
; APPLICANT: MUKUMOTO, Fujio
; APPLICANT: NISHIO, Shoichi
; APPLICANT: AKIMARU, Jiro
; APPLICANT: MITSUDA, Satoshi
; TITLE OF INVENTION: DNA Fragments Containing Biotin Biosynthetase Gene and
; TITLE OF INVENTION: Use of the Same
; FILE REFERENCE: 0152-0490P
; CURRENT APPLICATION NUMBER: US/09/180,109A
; CURRENT FILING DATE: 1998-12-03
; PRIOR APPLICATION NUMBER: 09/047838 JAPAN
; PRIOR FILING DATE: 1997-03-03
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Sphingomonas paucimobilis
; FEATURE:
; OTHER INFORMATION: Strain = JCM7511
; NAME/KEY: CDS
; LOCATION: (223)..(1245)
US-09-180-109A-23

Query Match 17.3%; Score 205.2; DB 3; Length 1420;
Best Local Similarity 52.4%; Pred. No. 5.7e-55;
Matches 509; Conservative 0; Mismatches 448; Indels 15; Gaps 2;

QY 184 TCAGTGAAGACACCACTCAACACCTGGACCAAGAAATTAAGCTATATATGACACA 243
DB 247 TCGAGGGGACCCCGGACCGACTGACCCCGCGGAGATCGCGCGCTGTTCGACCTG 306

QY 244 CCACTCATGGACTTGATGCACTATGCTCAGGTGCAACAGAGGTTCCAAACACTTCA 303
DB 307 CCCTTTCACGAGCTGTTGTTCCGCGCGGAGGTGACCCGCGCATCACCGCGCAT 366

QY 304 GAGGTTCAATTTGTGCACTTTATGAATATCAAACTGGTGGTTGTACCGAGCACTGTAAG 363
DB 367 CAGGTTCAAGTGTGACGCTGTTGTCGATCAAGACGGGCGGCTGCCCGAGGATTGCGGC 426

QY 364 TACTGTGCCAATCACAGCGTTTACAACTGTGTCAAGCTCAAGGCTGAAAGAAATCATCAAGTT 423
DB 427 TATTGAGGCACTGCAACCCATGCCGATACCGGGCTGAGGCGACCAAGCTGATGACCCG 486

QY 424 GATGAGGTGATTGAAGCTGCAAGAGGCAAGGCAATGGATCTTACAGGTTCTGTATG 483

Db 487 CGCGCCGTGTCAGAGGGCGCGCAGGCCAAGGATCAGGCTCGACGCGCTTCTGCATG 546
Qy 484 GGTGCTGCTTGGAGAGAGATGAAGGTAGAAAGTCAAACCTTGAAGAAAATCAAAGAGATG 543
Db 547 GCGCGGCTTGGCGCAACCCCAAGGATCG-----CGACATGCCCGCATCGTGGAGATG 600
Qy 544 ATCACTGCTGTCCATGACCTTGAATGGAGAGTTGTTCACCTCGGGAATGGTTGATAAA 603
Db 601 GTGAAGGGGTGCGCGCCATGGCCATGGAACCTGATGACGCTGGGCATGCTGACCGAT 660
Qy 604 GACCAAGCCACTGAATTGAAAGTGTGGTTCAGCGGTACACCCATCAACATTTGATACT 663
Db 661 GCACAGCGCAGACGCTCGCCGAGCGGCTGGACTATTACAAATCACAATATCGACAG 720
Qy 664 TAAAGGAACACTATCAAGGTGATCTCCACAAGAGCTTTTGATGATAGATTGAAAAACA 723
Db 721 TCGCCCGAGCGTTATGGCAGCGTCATCACCGCGCAGCTTCGGCGAGCGGTTGGAGAG 780
Qy 724 TTCAAAAAGTTCAAGGATCTGGATTAAAGGCATGCACAGGTGGTATTCTTGGTCTTGGT 783
Db 781 TTGGAGCATGTCCGCGATGCCGCGATCAATGTATGCTGTGGCGGTATTGTGCGCATGGT 840
Qy 784 GAGACCAAGAGACCGTGTATCTTCTCTACACCTTGGCCACATGGATCAGCATCCA 843
Db 841 GAGACGCGCGGACCGGTCGGCTTCATCCATGCGCTTGGCACCTCGCGGTCCATCCG 900
Qy 844 GAGTCTCTTCCAATCAACAGACTGTCCCAATCAAGGCGACGCCAATGTATGAAGAATT 903
Db 901 GGCAGCGTCCGGTGAACCGCTGGTCCGGTCAAGGCGACGGTATTGGGCGATATGTTG 960
Qy 904 AAGAACAGC-----AAGTTGAAGTTGATGAGATTGTGAGAACCATTCCTACTGCA 954
Db 961 GCCGACACGCCCGTGGCCAAAGATCGACGATATCGAATTCGTCCGCAACCGTCGCGTTGG 1020
Qy 955 AGATTGGTCAATCCCAACGCTCTATTATCAGATTGGCTGCAGGAGATATACAAATGAAGAG 1014
Db 1021 CGCATCACCATCCGCGATTTCGATGGTTCGCGCTGTCCGCGGGGCGGAGAGCATGTCGGAT 1080
Qy 1015 GCAGAACAGGTGATGTGCTTCATGGTGTGTTGTAATGCCATCTTACAGGTAAGAAAATG 1074
Db 1081 GCCACCCAGGCTTTGTGCTTCTGGGCGGCGGAACCTCGATCTTCAACGGCGCAAGCTG 1140
Qy 1075 CTCACAAACATGTGAACGGCTGGGATGAGGATAAAGCCATGTTGGCTAAATGGGGTCTG 1134
Db 1141 CTGACTGGGGCAATGCGGGCGACGACAAAGGACGAGCGCTCTTCGCCCGGCTGGGGCTC 1200
Qy 1135 AAACCAATGGAG 1146
Db 1201 ACGCCCATGGCG 1212

Search completed: September 17, 2005, 20:39:10
Job time : 251 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2005, 17:54:15 ; Search time 912 Seconds
(without alignments)
8678.876 Million cell updates/sec

Title: US-09-752-957F-1
Perfect score: 1188
Sequence: 1 atgctgtttattgactgc.....aggttcattcggtgcttga 1188

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333128559 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:**
1: /cgn2_6/ptodata/2/pubpna/PCT NEW PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US07 PUBCOMB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06 NEW PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06 PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07 NEW PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08 NEW PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08 PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09 PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09 NEW PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10 NEW PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11 NEW PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60 NEW PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60 PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1188	100.0	1188	21	US-09-752-957B-1
2	417.6	35.2	1125	21	US-10-741-849-6038
3	386.2	32.5	1334	18	US-10-425-114-28355
4	386.2	32.5	1345	20	US-10-425-115-180438
5	349.2	29.4	1659	9	US-09-740-288A-29
6	349.2	29.4	1659	16	US-10-431-544-29
7	328.2	27.6	3083	18	US-10-390-472-1

8	318.2	26.8	1420	19	US-10-437-963-18311	Sequence 18311, A
9	311.8	26.2	1515	9	US-09-740-288A-21	Sequence 21, Appl
10	311.8	26.2	1515	16	US-10-431-544-21	Sequence 21, Appl
11	310.2	26.1	1396	9	US-09-740-288A-17	Sequence 17, Appl
12	310.2	26.1	1396	16	US-10-431-544-17	Sequence 17, Appl
13	310.2	26.1	1439	9	US-09-740-288A-23	Sequence 23, Appl
14	310.2	26.1	1439	16	US-10-431-544-23	Sequence 23, Appl
15	310.2	26.1	1486	18	US-10-425-114-28968	Sequence 28968, A
16	310.2	26.1	1514	20	US-10-425-115-71987	Sequence 71987, A
17	306.6	25.8	1152	9	US-09-740-288A-7	Sequence 7, Appli
18	306.6	25.8	1152	16	US-10-431-544-7	Sequence 7, Appli
19	305.4	25.7	1443	19	US-10-767-701-13006	Sequence 13006, A
20	304	25.6	1137	9	US-09-938-842A-579	Sequence 579, App
21	304	25.6	1137	11	US-09-938-842A-579	Sequence 579, App
22	304	25.6	1137	21	US-10-702-364-1	Sequence 1, Appli
23	295.6	24.9	1477	9	US-09-740-288A-25	Sequence 25, Appl
24	295.6	24.9	1477	16	US-10-431-544-25	Sequence 25, Appl
25	294.4	24.8	1340	9	US-09-740-288A-11	Sequence 11, Appl
26	294.4	24.8	1340	16	US-10-431-544-11	Sequence 27, Appl
27	292.8	24.6	1526	9	US-09-740-288A-27	Sequence 27, Appl
28	292.8	24.6	1526	16	US-10-431-544-27	Sequence 27, Appl
29	292.2	24.6	1370	18	US-10-425-114-7490	Sequence 7490, Ap
30	292.2	24.6	1614	18	US-10-424-599-24740	Sequence 24740, A
31	279.8	23.6	1111	18	US-10-425-114-10388	Sequence 10388, A
32	260.8	22.0	1467	9	US-09-740-288A-19	Sequence 19, Appl
33	260.8	22.0	1467	16	US-10-431-544-19	Sequence 19, Appl
34	243.4	20.5	1032	9	US-09-740-288A-31	Sequence 31, Appl
35	243.4	20.5	1032	16	US-10-431-544-31	Sequence 31, Appl
36	237.6	20.0	987	17	US-10-282-122A-8411	Sequence 8411, Ap
37	228.6	19.2	957	17	US-10-282-122A-25092	Sequence 25092, A
38	223.6	18.8	1002	9	US-09-815-342-7073	Sequence 7073, Ap
39	223.6	18.8	1002	17	US-10-282-122A-22203	Sequence 22203, A
40	212.6	17.9	1038	17	US-10-282-122A-42017	Sequence 42017, A
41	211.6	17.8	1830121	17	US-10-329-670-1	Sequence 1, Appli
42	211.6	17.8	1830121	20	US-10-158-865-1	Sequence 1, Appli
43	211.6	17.8	1830121	22	US-10-981-687-1	Sequence 1, Appli
44	210.4	17.7	1907	18	US-10-425-114-14512	Sequence 14512, A
45	208.2	17.5	996	17	US-10-312-273-284	Sequence 284, App

ALIGNMENTS

RESULT 1

US-09-752-957B-1
; Sequence 1, Application US/09752957B
; Publication No. US20030104584A1
; GENERAL INFORMATION:
; APPLICANT: David SHUAN
; TITLE OF INVENTION: YEAST WITH HIGH BIOTIN PRODUCTIVITY AND THE PREPARATION METHOD 1
; FILE REFERENCE: C. utilis BIO2 - biotin synthase
; CURRENT APPLICATION NUMBER: US/09752,957B
; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Candida utilis
US-09-752-957B-1

Query Match 100.0%; Score 1188; DB 10; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGTCGTTTATTGACTGCTATTAGTCGTCGAGTCTCTTCCACTCTTAGAGTAGCT 60
Db	1	ATGTCGTTTATTGACTGCTATTAGTCGTCGAGTCTCTTCCACTCTTAGAGTAGCT 60
Qy	61	TCTAGGCTACTCTGGCAACAGGTGCTGCTGCTCGGAGATCTTGGAGATGTGTTTC 120
Db	61	TCTAGGCTACTCTGGCAACAGGTGCTGCTGCTGCTCGGAGATCTTGGAGATGTGTTTC 120


```
Db      816  ATTAGTTCCAAATCAGGGTACTCCAAATTTGTTGAAGTGTGAAAAATGCACCAAGGAAAG 875
Qy      912  GCAAGTTGAAGTTGATGAGATTGTGAGAACCATTTGCTACTGCAAGATTGGTCATGCCAAC 971
Db      876  ACAATTGGAATTTGATGCTATTTTGAGAACATTTGCTACTGCTAGATTGATTATGCTGTA 935
Qy      972  GTCTATTATCAGATTGGCTGCAGGAAGATATACAAATGAAAGAGGAGCAGAACAGGTGATGTG 1031
Db      936  ATCCATCATTAGATTGGCCCGGTCGATACACCATGAAGGAACACGAGCAGTTCTTTGTG 995
Qy      1032  CTTCAATGGCTGTTGTTAATGTCATCTTCAAGGTAAGAAATGCTCAACAATGTGTAA 1091
Db      996  TTTCAATGAGTGGGTCAATGCTATTTTACCGGTAAAGAGATGTTGACTACAATGTGTAA 1055
Qy      1092  CGCTGGGATCAGCATAAAGCCATGTTGGCTAAATGGGCTGTAACCAATGGAGAGTTT 1151
Db      1056  TGGTTGGGATGAAGATATCGCCATGTTGAAGAAATGGGTTTGAACCAATGGAAAGTTT 1115
Qy      1152  CA 1153
Db      1116  CA 1117
```

RESULT 3

US-10-425-114-28355
; Sequence 28355, Application US/10425114
; Publication No. US20040034889A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 28355

; LENGTH: 1334

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB4756-066-B10_FLI

US-10-425-114-28355

Query Match 32.5%; Score 386.2; DB 18; Length 1334;
Best Local Similarity 63.0%; Pred. No. 3.4e-106;
Matches 617; Conservative 0; Mismatches 353; Indels 9; Gaps 1;

```
Qy      169  TTGGAATGATGATTTGATGAGTGAAGACACCACTGACCTGACCAAGAAAGAAATTTAA 228
Db      170  TTTGAGAATGCACTCAATGCAAGACACCGGGCAAAATTTGACAAAGAGGAGATTTCA 229
Qy      229  GCTATATATGACACACCACTCATGGACTTGTATGACTATGCTCAGGTGCAACAGAGG 288
Db      230  GAGATATACACACTTCGCTGATTTGACCTTACCTATGCTTCGGCTTGGTCCACCGCGA 289
Qy      289  TTCCAAACAACCTTCAGAGTTCAATTTGCACTCTTATGAATATCAAACTGGTGGTTGT 348
Db      290  TTCCATGACCCCGTGCATTCAGATGTATACCTCTTCAACATCAAGACGGGGCTGTC 349
Qy      349  ACCGAGGACTGTAACTGTGTCCTCAATTCAGAGGTTACAACTGGTGTCAAGGCTGAA 408
Db      350  AGTGAAGACTGTTCTTACTGCGCCAGTCATCTCGTAGGACACCGGTCTCAAGGCACA 409
Qy      409  AGAATCATCCAAAGTTGATGAGTGAATGAAGCTGAAAGGAGGCAAGGCCAATGGATCT 468
Db      410  AAGCTCAGCTCTGTGCAATCGTCTCGAAGCTGCAATAGCTGCAATAGTAAAGAAACCGCAGC 469
```

Query Match 32.5%; Score 386.2; DB 20; Length 1345;

```
Qy      469  ACAAGGTTCTGTATGGTCTCTTGGAGAGAGATGAAAGGTAGAAAGTCAAACCTTCAAG 528
Db      470  TCAGATTTCTGCAATGGCGCGCGCTGGAGAGACATGAGGGGTCTGTAAGACGAACCTCAAG 529
Qy      529  AAAATCAAAGAGATGATCACTGCTGTCCATGACCTTGGAAATGAGAGAGTTGTGTCAACCTG 588
Db      530  AACATCAAAGAAATGATCAAGGGCTGCGGGCATGGGCATGAGGATGCGTGACCTT 589
Qy      589  GGAATGGTTGATAAAGACCAAGCCACTGAATGAAAGTCTCTGGTTGAGCGGTGTAACAC 648
Db      590  GGCATGGTCTGATCTCAAGCAAGAAAGAACTGAAAGATCGGGCTTGAAGGGCTTACAAT 649
Qy      649  CATAACATGATATCTTACAAGGAACACTATCCAAAGGTGATCTCCCAAGAAAGCTTTGAT 708
Db      650  CACAATGTGATGATGAGCGCGGAGCACTACCTAGCGTCAATCCACCAACGCTAGCTACGAC 709
Qy      709  GATAGATTGAAAAACATTTCAAAAAAGTTCAAGGATCTGGATTAAAGGATGACACAGGTGGT 768
Db      710  GAGGACTCAACACCATCAAGAAATGTCAGGAAGCAGGTATTACGTTGTCACCGCGGT 769
Qy      769  ATTCTTTGGTCTTGGTGAGACCCCAAGAGGACCGTGTATCTTTCTCTTACACTTGGCCACA 828
Db      770  ATTCTTTGGTCTGGGGGAAAAAGCCCGTGACCACTGGTCTCATCCACACTGTAGCTACA 829
Qy      829  ATGGATCAGCATCCAGAGTCTTTCCAATCAACAGACTGCTGCCATCAAGGGCACGCCCA 888
Db      830  CTACCCGCCCATCCAGAGTCTTTCCCGCTCAACGCCCTTGTCCCATCAAGAGGCACCCCA 889
Qy      889  ATGTATGAGAAGCTTAAGAACCAAGCAAGTTGAAGTTGATGAGATTGTGAGAACCATTTGCT 948
Db      890  CT-----AGGAGAAACCCCAAGCAATCTCTTTGATGCCATCTTCAGAACCATTTGCT 940
Qy      949  ACTGCAAGATTGGTTCATGCCAAAGCTCTATPATCAGATTGGCTGCAGGAAGATATACAATG 1008
Db      941  ACTGCAAGCTCTGCTGCTTCCACCACTCATCCGCTCTGCGAGCGCGCCACCATG 1000
Qy      1009  AAAGAGCAGAACAGGTGATGCTGCTCATGCTGGTGGTTGTAATGCCATCTTCACAGGTAAG 1068
Db      1001  CGTGAGGAAAAACAGATTCTGTGCTTCCAGGCTGGTGCAACCGCGTTCCTCTGCGGAG 1060
Qy      1069  ABAATGCTCAACAATGTGTAAGCTGGGATGAGATGAGATGAGATGAGTGGCTTAAATGG 1128
Db      1061  AGATGCTCAACACCGCGTGTATGTTGGTGGAAAGAGCAAGGCTATGTTTTCAGAGGTGG 1120
Qy      1129  GGTCTGAAACCAATGGAGA 1147
Db      1121  GGGCTTAGGCCAATGAAGA 1139
```

RESULT 4

US-10-425-115-180438

; Sequence 180438, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 180438

; LENGTH: 1345

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_9613C.1

US-10-425-115-180438

Best Local Similarity 63.0%; Pred. No. 3.4e-106;	
Matches 617; Conservative 0; Mismatches 353; Indels	9; Gaps 1;
Qy	169 TTGGAATATGCAATGTTCAGTGAAGACACACCACTCAACCTGGACCAAGAGAAATTTAA 228
Db	181 TTTGAGAAATGCACTCAATGCAAGGACCGCGGACAAATTGGACAAAGGAGGATTTCA 240
Qy	229 GCTATATATGACACACCACTCATGGAATTTGATGCACTATGCTCAGGTGCAACACAGAAAG 288
Db	241 GAGATATACAACACTTCGGCTGATTTGACCTTACCTATGCTTCGGGCTTTGGTCCACCGCGA 300
Qy	289 TTCCAAACAACTTCAGAGGTTCAATTTGTCACCTCTTATGATATATCAAACTGGTGGTTGT 348
Db	301 TTCCATGACCCCGCTGCCAATTCAGATGTGTACCTCTTTCAACATCAAGACGGCGGGCTGC 360
Qy	349 ACCGAGGACTGTAAAGTACTGTGCCCAATCAACAGCGTTTCAACACTGGTGTCAAGGCTGAA 408
Db	361 AGTGAAGACTGTTCCCTACTTCGGCCAGTCATCTCGTACGACACCGGTCTCAAGGCCACA 420
Qy	409 AGAATCATCCAAAGTTGATGAGGTGAATGAAGCTGAAAGGAGGCAAGGCCAATGTGATCT 468
Db	421 AAGCTCAGCTCTGTGCAATCCGTCCTCGAAGCTGCACGATAGCTAAGAAACCGGACG 480
Qy	469 ACAGGTTCTGTATGGGTGCTGCTTGAGAGAGATGAAAGTGAAGAGTCAAACTCAACTGAAG 528
Db	481 TCGAGATTCTGTCATGGGCGCGCTGGAGAGACATAGGGGTCTGTAGACGAACTCAAG 540
Qy	529 AAAATCAAAGAGATGATCACTGCTGTCATGACCTTGGAAATGGAGAGTTGTGTCAACGCTG 588
Db	541 AACATCAAGAAATGATCAAGGCGGTGGCGGCATGGGCATGGAGGCATGCGTACACCTT 600
Qy	589 GGAATGTTGTATAAGACCAAGCCACTGAATTTGAAAAAGTGTGGTGTGAAGCGGTACAAC 648
Db	601 GGCAATGTCGATGCTGCTCAAGCAAGGAACTGAAAGATGGCGGCTTGAACGCTTTACAAT 660
Qy	649 CATTAACATTGATACTTACAGGACACACTATCCAAAGGTGATCTCCACAAAGAGCTTTGAT 708
Db	661 CACAATGTCGATACGAGCGCGAGCATACCTTAGCGTCAATCAACACAGTACGTACGAC 720
Qy	709 GATAGATTGAAAAATTTCAAAAAAGTTCAAGGATCTGGATTAAAGGATGACACAGGTGGT 768
Db	721 GAGCGACTCAACACCATCAAGATGTCAGGAAGCAGGATTTACGTTGTCACCGCGGT 780
Qy	769 ATCTTGGTCTTTGGTGAAGCCAAAGAGACGGTGTATCTTTTCCTCTACACCTTGGCCACA 828
Db	781 ATCTTGGTCTGGGGGAAAAAGCCGTGACCCACGTCGCTCTCATCTCCACACTGTAGCTACA 840
Qy	829 ATGATCAGCATCAGAGTCTCTTCCCAATCAACAGATGGTGCCCAATCAAGGACACGCCA 888
Db	841 CTACCCGCCCATCCAGAGTCTCTCCCGTCAACGCCCTTGTCCCATCAAGAGGACACCCCA 900
Qy	889 ATGTATGAAGAAGTTTAAGAACCAAGTTGAAAGTTGATGAGATTGTTCAGAACCAATTGCT 948
Db	901 CT-----AGAGAAACCAAGCCATCTCTTTGATGCCATCTCTCAGAACCAATTGCT 951
Qy	949 ACTGCAAGATTGGTCAATGCCAAAGCTGTATTATCAGATTGGCTGCAGGAAGATATACAATG 1008
Db	952 ACTGCAGACTCGTGTCTCCACCAACCATCATCGCTCTCGCAGCGCGCGCCACACCATG 1011
Qy	1009 AAAGAGCGAAGCAGGTGATGTGCTTCATGGCTGGTTCTAATGCCATCTTTCAGAGGTAAG 1068
Db	1012 CGTGAGGAAAAACAGATTCTGTGCTTCCAGGCTGGTGCMAACGCCGCTTTCACGTGGCGAG 1071
Qy	1069 AAAATGCTCAACAAATGTGTAAACGGGTGGGATGAGGATAAAGCCATGTTGGCTCAAAATGG 1128
Db	1072 AAGATGCTCAACAGCGCGTGTATGTTGGTGGGAAGAACACAGGCTATGTTTCAGAGGTGG 1131
Qy	1129 GGTCTGAAACCAATGGAGA 1147
Db	1132 GGGCTTAGGCCAAATGAAGA 1150

RESULT 5

```

US-09-740-288A-29
; Sequence 29, Application US/09740288A
; Patent No. US2001003042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BH1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740, 288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 1659
; TYPE: DNA
; ORGANISM: Glycine max
US-09-740-288A-29

```

[illegible]

Db 685 GCTGGAATAACATATGCTCAGGTGGAATTAATTGGTCTTGGAGAGGCGAGAGGACCGT 744
Qy 802 GTATCTTTCTCTACACCTTGGCCACCAATGGATCAGCATCCAGAGTCTCTTCCCAATCAAC 861
Db 745 GTAGGATTTGGCATACTACCTGCTTCTGCTGACACCCAGAGAGTGTTCCTATTAAAT 804
Qy 862 AGACTGGTCCCAATCAAGGGCAGCGCAATGATGAGAGAGTTAAGAACCAAGTTGAA 921
Db 805 GCATTGGTTGCTGTCAGAGGCGCCCTTCAAG-----ATCAGAGCCCTGTAGAG 855
Qy 922 GTTCATGAGATTGTCAGAACCATCTGCTACTGCAAGATTGGTCATGCCAACTCTATTATC 981
Db 856 ATCTGGGAAATGATCCCGCATGATGTCACAGCTCGAATTTGTGATGCCAGAGCAATGGTG 915
Qy 982 AGATTGGCTGAGGAGAGATATACAAATGAAGAGGAGGAGGAGGAGGATGCTTCATGGCT 1041
Db 916 AGGCTTTTCAGAGAGAGAGTACGTTTCTCCATGCCAGAGAGCGCTGCTTCTTAGCT 975
Qy 1042 GGTGTAAATGCATCTTTCAGAGTAAGAAATGCTCACAACAATGTGTAAACGGCTGGGAT 1101
Db 976 GGGGCCAACTCCATCTTGGCTGGTGAGAAACTTCTTAAACAGCACAACAGCACTTTGAT 1035
Qy 1102 GAGATTAAGCCATTTGGCTTAATGGGCTCTGAAACCAATGG 1144
Db 1036 GCAGACCAGGCGATGTTCAAGATCCTTGGTCTGATTCCTCAAGG 1078

RESULT 9

US-09-740-288A-21
; Sequence 21, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740,288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Zea mays
US-09-740-288A-21

Query Match 26.2%; Score 311.8; DB 9; Length 1515;
Best Local Similarity 59.1%; Pred. No. 1.6e-83;
Matches 557; Conservative 0; Mismatches 377; Indels 9; Gaps 1;

Qy 202 AACACCTGGACCAAGAGAAATTAAGCTATATATGACACACCACTCATGGACTTGATG 261
Db 251 AACGACTGGAGCCGGCCGAGATCCAGCCGCTTACGACTACCGCTCTCGACTCTC 310
Qy 262 CACTATGCTCAGGTGCAACACAGAGGTTCCAAACCTTCAGAGGTTCAATTTGCACT 321
Db 311 TTTCACGGGGCTCAGGTCCACAGAAATGTCCATAAATTCAGAGAGTGCAGCAATGCACA 370
Qy 322 CTTATGATATCAAACTGGTGTGTACCGAGACTGTAGTACTGTGCCCAATCACAG 381
Db 371 CTTCTTTTCAATCAAGACTGGTGGATGTCAGTGAAGATTGTTTACTGTCTCAGTCATCA 430
Qy 382 CGTTACAACACTGGTGTCAAGCTTGAAGAAATCATCCAAGTTGTATGAGGTGATGAAGCT 441
Db 431 AGATACAACTGATTAAGAGCCCAAAATTTGATGACAAAGATGCTGTCTTGAAGCA 490
Qy 442 GCAAAGGAGGCAAGGCCAATGGATCTTACAAGGTTCTGTATGGGTGCTGTGGAGAGAG 501
Db 491 GCAAAAAGGCAAAAGAGTCTGGGAGCACCGCTTTTTCATGGGAGCTGTCATGGAGAGAA 550

Qy 502 ATGAAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGTCCATGAC 561
Db 551 ACCATTGGCAGGAATCAAACTTCAACAGATTCTTGAATATGTCAAGGAAATTAAGGGGT 610
Qy 562 CTTGGAATGAGAGTGTGTCACCTGGGAATGGTTCATAAAGACCAAGCCACTGAAATTG 621
Db 611 ATGGGCATGAGAGTGTGTCACACTAGGATGATAGAGAAACAACAGCTGGAAGACTC 670
Qy 622 AAAAGTCTGGTGTGACGGGTACAACTAATCATATTGATCTTTTCAAGGAAACATATCCA 681
Db 671 AAGAGSCTGACTTACAGCATATAATCATAACCTAGATACATCAAGAGAGTATTATCCC 730
Qy 682 AAGTGATCTCCACAGAGCTTTGATGATAGATTGAAACATTTCAAAAAGCTTCAAGGA 741
Db 731 AACATTATTACCAAGATCATATGATGATAGACTCTTTGAGCATGTCCGTGAA 790
Qy 742 TCTGGATTAAGGCATGCACAGGTGGTATTCTTGGTCTTGGTGAGACCCCAAGAGGACCGT 801
Db 791 GCTGGAATAAGCATCTGCTCAGGTGGAATCATTTGGTCTTGGTGAGAGAGGACCGG 850
Qy 802 GTATCTTTCTCTACACCTTGGGCACAAATGGATCAGCATCCAGAGTCTCTTCCCAATCAAC 861
Db 851 GTAGGTTGTGATACCTTAGCTACCTTGCCTACACACCCAGAGAGGTTCTCTATTAAAT 910
Qy 862 AGACTGGTCCCAATCAAGGCGACGCCAATGTATGAAGAGTTAAGAAACAAGCAAGTTGAA 921
Db 911 GCATTGGTGTGTTAAAGGCACACCTCTTTGAGG-----ACCAGAGCCCTGTAGAG 961
Qy 922 GTTGATGAGATTGTCAGAACCATTTGCTACTGCAAGATTGGTCATGCCCACTGTTATTATC 981
Db 962 ATCTGGGAAATGATCCGCATGTCGCCACTGCTCGGATCAGATGCCAAAGGCAATGGTG 1021
Qy 982 AGATTGGCTGCAGGAAGATATACAAATGAAGAGGACAGACAGAGTGATGTGCTTCATGGCT 1041
Db 1022 AGGCTTTTCAGAGCGCCGAGTACGGTCTTCGATGCCAGACAAGCGCTGTGCTTCTCGCT 1081
Qy 1042 GGTGTAAATGCCATTTTCACAGTAAGAAATGCTCACAACAATGTGTAAACGGCTGGAT 1101
Db 1082 GGGGCCAACTCCATCTTTGGCGGCGAGAAACTTCTCAACAACCCGCAACAGCACTTTGAT 1141
Qy 1102 GAGGATTAAGCCATTTGGCTTAATGGGCTCTGAAACCAATGG 1144
Db 1142 GCGAGCCAGGCGATGTTCAAGATCCTTGGCTGATCCCAAGG 1184

RESULT 10

US-10-431-544-21
; Sequence 21, Application US/10431544
; Publication No. US20030192073A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Kinney, Tony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/10/431,544
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Zea mays
US-10-431-544-21

Query Match 26.2%; Score 311.8; DB 16; Length 1515;
Best Local Similarity 59.1%; Pred. No. 1.6e-83;
Matches 557; Conservative 0; Mismatches 377; Indels 9; Gaps 1;

202 AACACCTGGACCAAGAGAAATTAAGCTATATATGACACACACCTCATGGACTTGATG 261
Db |||||
251 AACGACTGGAGCGCGCGAGATCCAGCGCTCTACGACTCACCGCTCTCGACCTCTC 310
Qy |||||
262 CACTATGCTCAGGTGCAACACAGAGGTTTCAACACCTTCAGAGGTTCAATTTGTCAC 321
Db |||||
311 TTTCACGGGGCTCAGGTCCACAGAAATGTCCATAAATTCAGAGAAGTGCAGCAATGCACA 370
Qy |||||
322 CTTTATGAATATAAAGCTGGTGTGACGAGGACTGTAAGTACTGTGCGCCCAATCACAG 381
Db |||||
371 CTTCTTTCAATCAAGACTGGTGTGATGATGAGGATGTTCTTACTGTCTCAGTCATCA 430
Qy |||||
382 CGTTACAACACTGGTGTCAAGGCTGAAAGAAATCATCCAAATGATGAGGTTGATGAAGCT 441
Db |||||
431 AGATACAACACTGGATTGAAGGCCCAAAATTTGATGAACAAAGATGCTGTCTTGGGAACA 490
Qy |||||
442 GCAAGGAGCAAGGCCAATGATCTACAAGGTTCTGTATGGTGTGCTTCGGAGAGAG 501
Db |||||
491 GCAAAAAGGCAAAAGAGTCTGGGAGCACCGGTTTTTGCATGGGAGCTGTCATGGAGAGAA 550
Qy |||||
502 ATGAAAGGTAGAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGTCTCCATGAC 561
Db |||||
551 ACCATTTGCGAGAAATCAAACTTCAACAGATCTTGAATATGTCAGAGAAATAAGGGGT 610
Qy |||||
562 CTTGGATGGAGAGTTGTGTCACTTGGGAATGGTGTGATGAAGCAAGCCACTGTAATG 621
Db |||||
611 ATGGCATGAGGTCTGTGACACACTAGGCATGATGAGAGAAACAAAGCTGAAGAACTC 670
Qy |||||
622 AAAAGTGTCTGGTTGACGGGTGACACCATTAACATTTGATTAATTTACAAGGAACATATCCA 681
Db |||||
671 AAGAAGGCTGGACTTACAGCATATAATCAATCAACCTAGATACATCAAGAGAGATATATCCC 730
Qy |||||
682 AAGGTGATCTCCACAAAGCTTTGATGATAGATGAAACATTTCAAAACAGTTCAAGGA 741
Db |||||
731 AACATTAATACCAAGATCATATGATGATGATGATGATGATGATGATGATGATGATGATG 790
Qy |||||
742 TCTGGATTAAGGCGATGACAGGTGTGTTCTTGGTCTTGGTGTGAGACCCCAAGAGGACCGT 801
Db |||||
791 GCTGGAATAAGCATCTGCTCAGGTGGAATCATTTGGTCTTGGTGAAGCAGAGGAGGACCG 850
Qy |||||
802 GTATCTTTCTTACACCTTGGCCCAATGATGATGATGATGATGATGATGATGATGATGATG 861
Db |||||
851 GTAGGGTGTGTGATACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 910
Qy |||||
862 AGACTGTCTCCAACTCAAGGCGACGCAATGTATGAAGAGTGAAGAACCAAGTGAAG 921
Db |||||
911 GCATTTGCTGTAAAGGCAACCTCTTGGG-----ACCAAGAGCTGTAGAG 961
Qy |||||
922 GTTATGAGATTTGTCAGAACCTTGTCTACGAAATGTTGTCATGTCAGCAAGCTTATATC 981
Db |||||
962 ATCTGGAAATGATCCGCATGATCGCCACTGCTCGGATCACGATGCCAAGGCAATGGTG 1021
Qy |||||
982 AGATTGCTCAGGAAGATATACATGAAGAGGCGAAGCAGGTGATGTCCTTCATGGCT 1041
Db |||||
1022 AGGCTTTTACAGGCGCGAGTACCGTTCTCGATGCCAGAACAGCGCTGTCTTCTCGCT 1081
Qy |||||
1042 GGTGTGATTCCTTCTTACAGGTGAAGAAATGCTCAACAATGTTGAACCGCTGGGAT 1101
Db |||||
1082 GGGGCCAACTCCATCTTTGCGGGGAGAACTTCTCAACCCGCAACCAACAGCTTTGAT 1141
Qy |||||
1102 GAGGATAAAGCCATTTGGCTTAAATGGGGTCTGAACCAATGG 1144
Db |||||
1142 GCGGACCGGCGATTTCAAGATCTTTGGCTGATCCCCAAGG 1184

RESULT 11
US-09-740-288A-17
; Sequence 17, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua

APPLICANT: Orozco, Emil
TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
FILE REFERENCE: BB1429 US NA
CURRENT APPLICATION NUMBER: US/09/740,288A
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/172929
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Microsoft Office 97
SEQ ID NO 17
LENGTH: 1396
TYPE: DNA
ORGANISM: Hordeum vulgare
US-09-740-288A-17
Query Match 26.1%; Score 310.2; DB 9; Length 1396;
Best Local Similarity 59.0%; Pred. No. 4.8e-83;
Matches 556; Conservative 0; Mismatches 378; Indels 9; Gaps 1;
Qy 202 AACACCTGGACCAAGAGAAATTAAGCTATATATGACACACCTCATGGACTTGATG 261
Db 213 AAGCACTGGACCGCGCGAGATCCAGGCCATCTACGACTCCCGCTCTCGACCTCTC 272
Qy 262 CACTATGCTCAGGTGCAACACAGAGGTTTCAACACCTTCAGAGGTTCAATTTGTCAC 321
Db 273 TTCCACGGGGCTCAAGTCCATAGGAATGTCATAAATTTAGAGAAGTGCACAATGCACA 332
Qy 322 CTTTATGAATATCAAACTGTGTGTCACGAGGACTGTAAGTACTGTGCGCCCAATCACAG 381
Db 333 CTTCTTTCAATTAAGACTGTGTGGTGCAGGAGATGTTCTATCTGCCCCACAGTCTTCA 392
Qy 382 CGTTACAACACTGGTGTCAAGGCTGAAGAAATCATCAAGTTGATGAGGTTGTAAGCT 441
Db 393 AGATACAGTACCGGATTGAAGCTGAAATAATTAATGAAGAAAGATGCCGCTCTAGAAGCA 452
Qy 442 GCAAGGAGGCAAGGCCAATGATCTACAAGGTTCTGTATGGTGTGCTGCTTGGAGAGAG 501
Db 453 GCTAAAGGCAAGAGGAGCTGGAGACCCGATTTTGCATGGGAGCCGATGGAGAGAG 512
Qy 502 ATGAAAGGTAGAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGTCTCCATGAC 561
Db 513 ACATTTGGCAGGAAACAACTTCAACAGATCTTGAATATGTCAGAGCATTAAGAGGT 572
Qy 562 CTTGGAATGGAGAGTTGTGTCACTTGGGAATGGTTGATTAAGACCAAGCCACTGTAATG 621
Db 573 ATGGGCATGGAGTCTGTGCACTTGGGCACTGTAGAGAAACAGCAAGCTGAAGAACTC 632
Qy 622 AAAAGTGTGGTGTGACGGCTGACACATTAACATTTGATGATGATGATGATGATGATGATG 681
Db 633 AAGAAGGCTGGAATTAACGCTTATAATCACTAGATACATCAAGAGAAATATTACCGG 692
Qy 682 AAGGTGATCTCCACAAAGAGCTTTGATGATAGATTGAAAACATTTCAAAAACGTTCAAGGA 741
Db 693 ACATTTATTTCTACAAGATCGTATGATGATGATGATGATGATGATGATGATGATGATG 752
Qy 742 TCTGGATTAAGGCAATGCAAGGTGATTTCTTGGTCTTGGTGTGAGACCCCAAGAGGACCGT 801
Db 753 GCTGGAATTAAGCGTCTGTCTCAGGTGGAATTAATTTGGTCTTGGAGAGCGGAGAGACCGT 812
Qy 802 GTATCTTTCTCTACACCTTGGCCCAATGATGATGATGATGATGATGATGATGATGATGATG 861
Db 813 GTAGGCTGTGTGCATACACTGCGCCATTTTGGCAACACACCCACAGAGAGTGTTCCTATCAAT 872
Qy 862 AGACTGTCTCCCAATCAAGGCGCACGCAATGATGAAGAAAGTTTAAGAAACCAAGCAAGTTGAA 921
Db 873 GCATTTGCTGTCAAGGCGACCGCTCTTCAAG-----ATCAGAGAGCTGTAGAG 923
Qy 922 GTTGTAGATGATGTCAGAACCTTGTCTACTGCAAGATGTTGGTCAATGCCAACGCTTATATC 981
Db 924 ATATGGGAATGATCGCATGATTTGCCAGCGCTCGGATTTGTGATGATGATGATGATGATG 983
Qy 982 AGATTGGCTGCGAGGAAGATATACATGAAGAGGCAAGAGTGTGCTTCATGGCT 1041

Db 984 AGACTTTCCGAGGCGAGTACGGTTCTCCATGCCAGAGCAAGCTCTCTGCTTCTGCT 1043
Qy 1042 GGTGTAAATCCATCTTTACAGAGTAAGAAAATGCTCAACAATGTGTAAAGCGCTGGGAT 1101
Db 1044 GGGCCAACTCGATCTTTGCGCGTGAAAGCTCTCGACAACTGCAAAACAGGACTTTGAT 1103
Qy 1102 GAGGATAAGCCATGTTGGCTTAATGGGCTGTAACCAATGG 1144
Db 1104 GCGGACCAGGCAATGTTCAAGATCTTGGCTGATTCCTCCCAAGG 1146

RESULT 12
US-10-431-544-17
; Sequence 17, Application US/10431544
; Publication No. US20030192073A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Kinney, Tony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/10/431.544
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Hordeum vulgare
US-10-431-544-17

Query Match 26.1%; Score 310.2; DB 16; Length 1396;
Best Local Similarity 59.0%; Pred. No. 4.8e-83;
Matches 556; Conservative 0; Mismatches 378; Indels 9; Gaps 1;

Qy 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTCATGGACTTGATG 261
Db 213 AACGACTGGACCGCCGCGAGATCCAGGCCATCTAGACTCCCGCTCTCGACCTCTC 272
Qy 262 CACTATGCTCAGGTGCAACACAGAGGTTCCAACTTCAGAGTTCAATTTGAGT 321
Db 273 TTCCACGGGGCTCAAGTCCATAGGAATGTCATTAATTTAGAGAGTGCAACAATGCAC 332
Qy 322 CTTTATGAATCAAACTGTTGTGTACCGAGGACTGTAAAGTACTGTGCCCAATCACAG 381
Db 333 CTTCTTTCAATAAGAGCTGTGGGTGCGAGAGATTTTCATCTGCCCCACAGTCTTCA 392
Qy 382 CGTTACAACACTGCTGTCAAGGCTGAAAGAAATCATCCAGTTGATGAGGTGATTGAAGCT 441
Db 393 AGATACAGTACCGGATTTGAAGCTGAAATTAATGAAGAAAGATGCCGTCTTAGAAGCA 452
Qy 442 GCAAGAGGCAAGGCAATGATCTCAAGGTTCTGTATGGTGTGCTTGGAGAGAG 501
Db 453 GCTTAAAGGCAAGGAGGCTGGGAGCACCCGATTTTGTATGGGAGCGCATGGAGAGAG 512
Qy 502 ATGAAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGTGTCATGAC 561
Db 513 ACAATTGGCAGGAAACAAACTTTCAACAGATTTCTTGAATATGTCAAGGACATGAAGGT 572
Qy 562 CTTGGAATGGAGAGTTGTGTCAACCTGGGAATGGTTGATAAAGCAACCAAGCACTGAATTG 621
Db 573 ATGGGATGGAGTCTGTTGCCACCTGGGATCTGTAGAGAAACAGCAAGCTGAAGAACTC 632
Qy 622 AAAAGTGTGGTTGACGGGTACAACTTCAATGATATTTTCAAGGAACATATCCA 681
Db 633 AAGAAGGCTGGACTTACAGCTTATATCAATCAACCTAGATACATCAAGAGAAATATTACCCG 692
Qy 682 AAGGTATCTCCACAGAGAGCTTTGATGATAGATTGAACATTTCAAAACAGTTCAAGA 741
Db 693 AACATTTATTTCTACAAGATCGTATGATGATGATTAAGAAATATGATGATGCTCTTGGAGCA 752

Qy 742 TCTGGATTAAGGCATGCAAGGTGTATTTCTTGGTCTTGGTGAGACCCCAAGAGGACCGT 801
Db 753 GCTGGAATAAGCGCTCTGCTCAGGTGGAATTTATTTGGTCTTGGAGAGCGGAGGAACCGT 812
Qy 802 GTATCTTTCTCTACACCTTGGCCCAATAGGATCAGCATCCAGAGTCTCTTCCAATCAAC 861
Db 813 GTAGGGCTGTGTGATACACTGGCCACTTTGGCCAAACACACCCAGAGAGTGTTCCTATCAAT 872
Qy 862 AGACTGCTCCCAATCAAGGCGACGCCAATGTATGAAGAAGTTAAGAACCAAGCAAGTTGAA 921
Db 873 GCATTGATTGCTCTCAAGGCGACGCTCTTCAGG-----ATCAGAAGCCTGTAGAG 923
Qy 922 GTTGATGAGATTGTGAGAACCATTTGCTACTGTCAAGATTGGTCAAGCAAGCTCTATTATC 981
Db 924 ATATGGGAAATGATCCGCATGATTGCCAGCGCTCGGATTTGTGATGCCAAAGGAATGGTG 983
Qy 982 AGATTGGCTGCGAGGAAGATATCAATGAAGAAGGCGAGACAGGTGATGTCTTTCATGGCT 1041
Db 984 AGACTTTCCGCGAGGCGAGTACGGTTCTCCATGCCAGAGCAAGCTCTCTGCTTTCTTGT 1043
Qy 1042 GGTGTAATGCCATCTTTCACAGGTAAGAAAATGCTCAACAATGTGTAAACGGCTGGGAT 1101
Db 1044 GGGGCCAACTCGATCTTCCGCGGTGAAGAGCTCTGCAACTGCCAAACACAGACTTTGAT 1103
Qy 1102 GAGGATAAAGCCATGTTGGCTAAATGGGCTGTGAACCAATGG 1144
Db 1104 GCGGACCAGGCAATGTTCAAGATCTTGGCTGATTCCTCCCAAGG 1146

RESULT 13
US-09-740-288A-23
; Sequence 23, Application US/09740288A
; Patent No. US20010039042A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; TITLE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740, 288A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1439
; TYPE: DNA
; ORGANISM: Zea mays
US-09-740-288A-23

Query Match 26.1%; Score 310.2; DB 9; Length 1439;
Best Local Similarity 59.0%; Pred. No. 4.8e-83;
Matches 556; Conservative 0; Mismatches 378; Indels 9; Gaps 1;

Qy 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTCATGGACTTGATG 261
Db 235 AACGACTGGAGCGGCGCGAGATCCAGGCGTCTAGCATCTCCGCTCTCGACCTCTC 294
Qy 262 CACTATGCTCAGGTGCAACACAGAGAGTTTCCAAACCTTTCAGAGGTTCAATTTGTGCACT 321
Db 295 TTTCACGGGCTCAGTCCACAGAAATGTCCATAAATTCAGAGAGTGCAGCAATGCACA 354
Qy 322 CTTATGAATATCAAACTGTTGTATCCGAGGAGCTGTAAAGTACTGTGCCCAATCACAG 381
Db 355 CTTCTTTCAATCAAGACTGGTGGATGCGAGTGAAGATTGTTTACTGCTCTCAGTCA 414
Qy 382 CGTTACAACACTGTTGTCAGGCTCAAGAAATCAATCCCAAGTTGATGAGGTGATTGAAGCT 441
Db 415 AGATACAACACTGGATTGAAGGCCCAAAATATGATGAACAAATATGCTGCTTGGAGCA 474

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 28968
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4767-018-H9_FLI
US-10-425-114-28968

Query Match 26.1%; Score 310.2; DB 18; Length 1486;
Best Local Similarity 59.0%; Pred. No. 4.9e-83;
Matches 556; Conservative 0; Mismatches 378; Indels 9; Gaps 1;

Qy 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACCACTCATGGACTTGATG 261
Db |||||
Qy 241 AACGACTGGAGCGCGCCGAGATCCAGGCCGTCTACGACTCACCGCTCTCGACCTCCTC 300
Db |||||
Qy 262 CACTATGCTCAGGTGCAACACAGAGGTTCCAAACACCTTCAGAGGTTCAATTGTGCACT 321
Db |||||
Qy 301 TTTCACGGGGCTCAGGTCCACAGAAATGTCCATAAATTCAGAGAAAGTCAGCAATGCACA 360
Db |||||
Qy 322 CTTATGAATATCAAACTGGTGTGTACCGAGACTGTAAAGTACTGCGCCCAATCACAG 381
Db |||||
Qy 361 CTTCCTTCAATCAAGCTGGTGGATGCACTGAAGATGTTCTTACTGCTCAGTCATCA 420
Db |||||
Qy 382 CGTTACAACTGGTGTCAAGGCTGAAAGAAATCAACAGATGATCACTGCTGTCCATGAC 441
Db |||||
Qy 421 AGATACAACTGGATTTGAAGGCCCAAAATTTGATGAACAAATATGCTGTCTTGAAGCA 480
Db |||||
Qy 442 GCAAGAGGCAAGGCCAATGGATCTCAAGGTTCTGTATGGTGTCTGTGGAGAGAG 501
Db |||||
Qy 481 GCAAAAAGGCAAAAGAGTCTGGAGCACCCGTTTTCATGGGAGCTGCGATGGAGAA 540
Db |||||
Qy 502 ATGAAGGTAGAAAGTCAAACTCAAGAAATCAAGAGATGATCACTGCTGTCCATGAC 561
Db |||||
Qy 541 ACCATTGGCAGGAATCAAACTCAACAGATCTTGAATATGTCAAGGAAATTAAGGGGT 600
Db |||||
Qy 562 CTTGGAATGGAGAGTTGTGTCACTGGGAATGGTTGATTAAGACCAAGCCACTGAAATTG 621
Db |||||
Qy 601 ATGGCATGGAGGTCTGTTCACACTAGGCATGATAGAGAAACAAAGCTGAAGACTC 660
Db |||||
Qy 622 AAGAGTCTGGTTGACGGGTACAACTAACTTGAATCTTACAGGAACTATATCCA 681
Db |||||
Qy 661 AAGAAGGCTGGACTTACAGCATATAATCAAACTAGATACATCAAGAGAGTATATCCC 720
Db |||||
Qy 682 AAGTGATCTCCACAGAGCTTTGATGATAGATTGAATCAAAACGTTCAAGGA 741
Db |||||
Qy 721 AACATTATACCAAGATCATATGATGATAGACTGTGAGCATCTTGAAGCATGTCGGTGA 780
Db |||||
Qy 742 TCTGGATTAAGGCATGCACAGGTGGTATTTCTGGTCTTGGTGAGACCCCAAGAGACCGT 801
Db |||||
Qy 781 GCTGGAATAGCATCTGCTCAGGTGGAATCATTTGGTCTTGGTGAAGCAGAGGAGCCGG 840
Db |||||
Qy 802 GTATCTTTCTTACCTTTGGCCACAATGGAATCAGATCAAGATCTCTTTCCAAATCAAC 861
Db |||||
Qy 841 GTAGGGTTGTTGCATACCTTAGCTACCTTGCCTACACACAGAGAGCGTTCTTATTAAT 900
Db |||||
Qy 862 AGACTGCTCCCAATCAAGGCAAGCCCAATGATGAAGAGTTAAGACACAGCAAGTTGAA 921
Db |||||
Qy 901 GCATTGGTTGCTGTAAGAGGCAACCTCTTTGAGG-----ACCAAGAGCCTGTAGAG 951
Db |||||
Qy 922 GTTGATGAGATTGTCAGAACCAATTGCTACTGCAAGATTGGTCATGCCAACGTCATTATC 981
Db |||||

Search completed: September 17, 2005, 21:28:17
Job time : 920 secs

Db 952 ATCTGGAAATGATCCGCATGATGCCCACTCTCGGATCAGCATGCCAAAGCAATGGTG 1011
Qy 982 AGATTGGCTGCAGGAAGATATACAAATGAAGAGGCGAGAACAGGTGATGCTTTCATGGCT 1041
Db 1012 AGGCTTTTCAGCAGGCCGAGTACGGTTCTCGATGCCAGAACAGGCGCTGTGCTTCTCGCT 1071
Qy 1042 GGTGTAATGCTCTTTCACAGGTAAGAAATGCTCAACAATGTGTAACGGCTGGGAT 1101
Db 1072 GGGGCCAACTCCATCTTTGCGGCGAGAACTTCTCAACCGCAACACGACTTTGAT 1131
Qy 1102 GAGGATAAAGCCATGTTGGCTTAAATGGGGTCTGAACCAATGG 1144
Db 1132 GCGGACCAGGCGATGTTCAAGATCCTTTGGCCTGATCCCAAGG 1174

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2005, 15:54:19 ; Search time 4431 Seconds
(without alignments)
10205.451 Million cell updates/sec

Title: US-09-752-957F-1

Perfect score: 1188

Sequence: 1 atgtcggttatattgactgc.....aggtgcattcggtgttga 1188

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	333.4	28.1	969	CNS06HEG	AL398894 T7 end of
2	330.4	27.8	985	CNS06234	AL421814 T3 end of
3	297	25.0	1301	CNS0A9XF	AX819210 Arabidops
4	296.2	24.9	955	CNS07350	AL427092 clone BA0
5	294.4	24.8	1306	CNS0A9QQ	AX818983 Arabidops
6	292	24.6	1412	CY109523	AY109523 Zea mays
7	291	24.5	969	CO004857	CO004857 EST793192
8	290.8	24.5	876	CO084519	CO084519 GR_Ea01D
9	270	22.7	872	CF822545	CF822545 EST69927
10	266	22.4	834	CF822544	CF822544 EST69926
11	262.4	22.1	829	CF7219376	CF7219376 CCABF1670
12	252.4	21.2	862	CF719376	CF719376 CCABF1670
13	251.6	21.2	781	CD852378	CD852378 DH0ALL272
14	251.4	21.2	770	BU063672	BU063672 Fgr_3 G19
15	250.2	21.1	671	CAY81203	CAY81203 017Fo7AF
16	248.6	20.9	780	CF680028	CF680028 CCAD710TF
17	242	20.4	729	BU059880	BU059880 Fgr-C 0 G
18	240.6	20.3	811	CF448691	CF448691 EST685036
19	239.4	20.2	865	CF682137	CF682137 CCAGC25TR
20	239.4	20.2	948	CF721975	CF721975 CCABF1670
21	239.2	20.1	635	BU634775	BU634775 008D07 In
22	237.2	20.0	822	CO147058	CO147058 EST822111
23	237	20.0	825	CF690143	CF690143 CCAB241TF
24	234.4	19.7	721	BU064578	BU064578 Fgr_5_D19

25	233.4	19.6	509	8	AZ931212	AZ931212 474.dhz63
26	233	19.6	644	1	AV9233483	AV9233483 AV9233483
27	232.4	19.6	694	6	CA069980	CA069980 SCSGAD100
28	231.4	19.5	703	5	BU062717	BU062717 Fgr_2_C13
29	229.2	19.3	646	5	BU062716	BU062716 Fgr_2_C13
30	228.6	19.2	975	9	CNS06E08	AL394500 T3_end_of
31	227.4	19.1	596	1	AV9233288	AV9233288 AV9233288
32	226.6	19.1	730	1	AI960629	AI960629 sc87b11.Y
33	225.2	19.0	707	7	CF682132	CF682132 CCAGC25TF
34	225.2	19.0	826	7	CF719378	CF719378 CCAEF16TR
35	219.8	18.5	891	7	CO117806	CO117806 GR_EB01N
36	217.2	18.3	585	1	AI775752	AI775752 EST256852
37	216.2	18.2	730	5	BU025073	BU025073 QHF7113.Y
38	206.6	17.4	766	7	CN145395	CN145395 WOUND1_28
39	206	17.3	724	6	CA583129	CA583129 EST002804
40	204.2	17.2	796	4	BG582087	BG582087 EST483825
41	204	17.2	831	6	CD380074	CD380074 PTM05243
42	203.2	17.1	604	7	CO136883	CO136883 EST831554
43	201.2	16.9	588	2	AW980387	AW980387 EST391540
44	199.4	16.8	681	7	CF449199	CF449199 EST68544
45	199.2	16.8	814	6	CB293284	CB293284 UCRCS01_0

ALIGNMENTS

CNS06HEG 969 bp DNA linear GSS 30-NOV-2001
T7 end of clone AS0AA011H02 of library AS0AA from strain CLIB 533
of Saccharomyces bayanus, genomic survey sequence.

AL398894
AL398894.1 GI:12152774

GSS.

SOURCE

ORGANISM

Saccharomyces bayanus

Saccharomyces bayanus

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 969)

AUTHORS

Souciot,J.L., Aigle,M., Artiguenave,F., Blandin,G.,

Botolin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,

de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,

Malperuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,

Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,

Wincker,P. and Weissenbach,J.

Genomic exploration of the hemiascomycetous yeasts: 1. A set of

yeast species for molecular evolution studies

FEBS Lett. 487 (1), 3-12 (2000)

20584711

REFERENCE 2 (bases 1 to 969)

AUTHORS

Bon,E., Neuveglise,C., Casaregola,S., Artiguenave,F., Wincker,P.,

Aigle,M. and Durrens,P.

Genomic exploration of the hemiascomycetous yeasts: 5.

Saccharomycetes bayanus var. uvarum

FEBS Lett. 487 (1), 37-41 (2000)

20584715

REFERENCE 3 (bases 1 to 969)

AUTHORS

Genoscope.

Direct Submission

Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,

2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :

secref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

This GSS is part of a random genomic sequencing program of thirteen

yeast species: Saccharomyces bayanus var. uvarum, Saccharomycetes

exiguus, Saccharomycetes servazii, Zygosaccharomyces rouxii,

Saccharomycetes kluyveri, Kluyveromyces thermotolerans, Kluyveromyces

lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia

angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,

Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to

5 kb were prepared and both extremities were sequenced. See

keywords for description of this sequence and for the sequence of

```

the other extremity of this insert.
FEATURES
    source
        1..969
            /organism="Saccharomyces bayanus"
            /mol_type="genomic DNA"
            /strain="CLIB 533"
            /variety="uvaxum"
            /db_xref="taxon:4931"
            /clone="AS0A011H02"
            /clone_lib="AS0A"
            /note="end : 17"
            <2..>658
            /note="similar to Saccharomyces cerevisiae ORF YGR286c [
            BIO2 ; biotin synthetase ]"
            /evidence=not_experimental

misc_feature
    28.1%; Score 333.4; DB 9; Length 969;
    Best Local Similarity 71.0%; Pred. NO. 4.4e-83;
    Matches 458; Conservative 1; Mismatches 177; Indels 9; Gaps 1;

ORIGIN
Query Match      28.1%; Score 333.4; DB 9; Length 969;
Best Local Similarity 71.0%; Pred. NO. 4.4e-83;
Matches 458; Conservative 1; Mismatches 177; Indels 9; Gaps 1;

Qy 522 CTTGAAGAAATCAAGAGATGATCAGCTGCTCCATGACCTTGGATGGAGAGTTGTGT 581
Db 4 CATGAAGAGAAATGAGGAGATGATTTCAAAGTGAACGACATGGGGCTGGAAACATGTGT 63

Qy 582 CACCTGGGATGTTGATAAGACCAAGCCACTGAATTTGAAAGTGTCTGGTTGACGGC 641
Db 64 TACTTTGGTATGTCGACCAAGACCAGCAAGCAGCTAAAGATGACAGTGTGACCGC 123

Qy 642 GTACAACCATTAACATTGATCTTCAAGGAACACTATCAAAAGGTGATCTCCACAAGAG 701
Db 124 ATACAACCATTAACATTGATCTTCAAGGAACACTATCAAGTAAGTCAATCAACCCAGAAC 183

Qy 702 CTTTGATGATGATTAAGAAACATTCAAAAAGTTCAAGGATCTGATTAAGGATGACAC 761
Db 184 CTATGACATGAGCTGCAGACCATTTAAGACGTTCCAAAGAGTCCGGRATANAAGCTGCAC 243

Qy 762 AGGTGGTATTCTGTGTTGTTGACCAAGAGGACCGTGTATCTTTCCCTCTACACCTT 821
Db 244 TGGTGGTATCTTAGTCTAGGTGAAGCGAAGATGACCAATGTTGTTTCATTATACAGTT 303

Qy 822 GGCCACAATGGATCAGCATCCAGAGTCTTCTCCAAATCAACAGACTGGTCCCAATCAAGG 881
Db 304 GTCCAAACATGACCCACATCCAGAGTCCCTACCAATCAACAGACTGTTGTCTATCAAGG 363

Qy 882 CACGCCATGATGAGAGAGTTAAGAAC-----AAGCAAGTTGAAGTTGATGATGAT 932
Db 364 GACTCCAAATGGCTGAGGAAGTTCGCCACCGCAAGTAAGAAAGTTGCAATTCGATGAAT 423

Qy 933 TGTCGAACCATTTGCTACTGCAAGATTTGGTTCATGCCAACGTTCTATTATCAGATTGGCTGC 992
Db 424 CTTAAGAACTATCGCCACAGCAGCAATTTGTCATGCTAAGCGGATCATAGACTTGTCTGC 483

Qy 993 AGGAAGATATCAATGAAGAGGCGAGAACAGGTGATGTCTTCATGCTGTTGTTGTAATGC 1052
Db 484 TGGTGGTTACAAATGAAGGAGACCGAACAAATTTGCTGCTTCATGCTGTTGTTGTAACAG 543

Qy 1053 CATCTTCAGGTAGAAAATGCTCACAACATGTTAAGCGTGGATGAGGATTAAGC 1112
Db 544 TATTTTCACCGGTAGAAAATGCTTACAACATGTTGTAACGGTGGGATGAAGCAAGC 603

Qy 1113 CATGTTGGCTAAATGGGGTCTGAAACCAATGGAGAGTTTCAAATA 1157
Db 604 AATGTTGGCTAAGTGGGGCTTGCAACCTATGGAAGCCTTCGATA 648

RESULT 2
CNS06234/c
LOCUS
DEFINITION
    985 bp DNA linear GSS 06-JUL-2001
    T3 end of clone AY0AA015C04 of library AY0AA from strain CBS 6340
    of Kluyveromyces thermotolerans, genomic survey sequence.
ACCESSION
    AL421814
VERSION
    AL421814.1 GI:12205009

```

```

KEYWORDS
SOURCE
    ORGANISM
        Kluyveromyces thermotolerans
        Kluyveromyces thermotolerans
        Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
        Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE
    1 (bases 1 to 985)
        Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
        Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
        de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
        Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
        Saurin,W., Tekala,P., Toffano-Nioche,C., Wesolowski-Louvel,M.,
        Wincker,P. and Weissenbach,J
        Genomic exploration of the hemiascomycetous yeasts: 1. A set of
        yeast species for molecular evolution studies
        FEBS Lett. 487 (1), 3-12 (2000)
        20584771
        11152876
        2 (bases 1 to 985)
        Malpertuy,A., Llorente,B., Blandin,G., Artiguenave,F., Wincker,P.
        and Dujon,B.
        Genomic exploration of the hemiascomycetous yeasts: 10.
        Kluyveromyces thermotolerans
        FEBS Lett. 487 (1), 61-65 (2000)
        20584720
        11152885
        3 (bases 1 to 985)
        Genoscope.
        Direct Submission
        Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
        2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. [E-mail :
        seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]
        This GSS is part of a random genomic sequencing program of thirteen
        yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
        exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
        Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
        lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
        angusta, Debaryomyces hansenii var. Hansenii, Pichia sorbitophila,
        Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
        5 kb were prepared and both extremities were sequenced. See
        keywords for description of this sequence and for the sequence of
        the other extremity of this insert.
FEATURES
    source
        1..985
            /organism="Kluyveromyces thermotolerans"
            /mol_type="genomic DNA"
            /strain="CBS 6340"
            /db_xref="taxon:4916"
            /clone="AY0AA015C04"
            /clone_lib="AY0AA"
            /note="end : T3"
            complement(<1..>788)
            /note="similar to Saccharomyces cerevisiae ORF YGR286c [
            BIO2 ; biotin synthetase ]
            1 putative frameshift(s)"
            /evidence=not_experimental

misc_feature
    27.8%; Score 330.4; DB 9; Length 985;
    Best Local Similarity 66.0%; Pred. No. 3.2e-82;
    Matches 492; Conservative 1; Mismatches 252; Indels 1; Gaps 1;

ORIGIN
Query Match      27.8%; Score 330.4; DB 9; Length 985;
Best Local Similarity 66.0%; Pred. No. 3.2e-82;
Matches 492; Conservative 1; Mismatches 252; Indels 1; Gaps 1;

Qy 167 CATTGGAATATGCAATTTGCTAGTGAAGACACACAGTCAACCTGGACCAAGAAATTA 226
Db 796 CACTGCAGTTTCCGSSGTGCTCGCTCGAGCAGCCTTCGCCAGAGTGGTCCAAAGAGCAGCTGC 737

Qy 227 AAGCTATATATGACACACCACTCATGGACTTGATGCACTATGCTCAGGTCAACACAGAA 286
Db 736 GAGAGATCTACACACGCCGCTGCTGGAGCTTGTGCACCTCGGCCAGCTGCAGACCGCC 677

Qy 287 GGTTCCAACACCTTCAGAGGTTCAATTTGTCACCTCTTATGAATATCAAAACTGGTGT 346
Db 676 GCTGGCAGATCTCTACCAAGTGCAGCTGTGCACTCTTATGAACATCAAGAGCGGTGGT 617

```

```
QY 347 GTACCGAGGACTGTAAAGTACTGTGCCAATCAAGCGTTTCAACACTGGTGTCAAGGCTG 406
Db 616 GTTCCCGAGGACTGTAAAGTACTGTGCCGAGTCTGCGCGGTACTCGAGGGGCGTCAAGGCGG 557
QY 407 AAGAATATCAATCCAAAGTTGATGAGTGTGATGAAGCTGCAAAAGGAGGCAAGGCGCAATGGAT 466
Db 556 AGAAGCTGTATCCGTGATGCGGTGCTCAAGAGGCGGAGGAGGCGCAAGCAGACGAT 497
QY 467 CTACAAGGTTCTGTATGGGTGCTGTTGGAGAGAGATGAAGGTAGAAAATGTTCAAACTTGA 526
Db 496 CCACTGCTTTCGCTGGCGCGGCTGCGGTGACATGAACGGGCGTAAAGTCCGGS-TGC 438
QY 527 AGAATATCAAGAGATGATCACTGCTGTCCATGACCTTGGATGGAGAGTGTGTCAACC 586
Db 437 GCAAGATCTCCGAGATGGTCAAGAAGTCAATGAATGGGCGCTCGAGAGCGTGTGTACGC 378
QY 587 TGGGAATGTTGATAAAGACCAAGCCACTGAATTTGAAAGTGTGCTGGTGTGACGGCGTACA 646
Db 377 TGGGCATGTCAAGGAGTACGCGAGAGCTGAAGAGCGCGGAGCTCACCGCGTACA 318
QY 647 ACCATAAATGATCTTCAAGGAACACTATCTCAAGGTGATCTCCACAAGAGCTTTG 706
Db 317 ACCATAAATGATCTTCAAGGAACACTATCTCAAGGTGATCTCCACAAGAGCTACG 258
QY 707 ATGATAGATGAAAACATTTCAAAAAGCTTCAAGGATCTGGATTAAGAGGATGACAGGTG 766
Db 257 AGCAGCGTGTGAACAGATCAAGAACGCTCAAAAGCGCGGCGCATCAAGGGGTGACCGGTG 198
QY 767 GTATTTCTGCTGTGGTGAGACCAAGAGGACCGTGTATCTTTCCTACACTGGCGCA 826
Db 197 GTATTTCTGCGGTGGCGGAGACAGAGGAGGACCAATGGTGTGCTTCCACACTTGGCCA 138
QY 827 CAATGATCAGCATCAGAGTCTCTTCCCAATCAACAGACTGGTCCCAATCAAGGGCAGCG 886
Db 137 ACATGACACACACCGGAGTGGTCCCAATCAACAGACTGGTCCCAATCAAGGGACCC 78
QY 887 CAATGTATGAAGAAGTTAAGAACAG 912
Db 77 CCATGGAGGAGGATCTGCAGAAAAAG 52
```

```
RESULT 3
CNSOAX9F 1301 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTFB57ZG05 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
ACCESSION BX819210
VERSION BX819210.1 GI:42469429
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1301)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1301)
GENOSCOPE Genoscope.
DIRECT SUBMISSION Direct Submission
AUTHORS Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
```

```
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGU INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
Source 1..1301
/organism="Arabidopsis thaliana"
/mol_type="mrna"
/db_xref="taxon:3702"
/clone="GSLTFB57ZG05"
/tissue_type="Flowers and buds"
/plasmid="pCMWSPORT 6"
complement(1..1301)
/gene="At2g43360"
gene
ORIGIN
Query Match 25.0%; Score 297; DB 3; Length 1301;
Best Local Similarity 58.7%; Pred. No. 1.1e-72;
Matches 538; Conservative 0; Mismatches 370; Indels 9; Gaps 1;
QY 202 AACACTGACCAAGAAAGAAATTAAGCTATATATGACACACCACTCATGACTTGATG 261
Db 245 AACGATGGAGTAGAGATGAATCAAGCTGTTTATGATTCCTCTTGTGACCTCTC 304
QY 262 CACTATGCTAGGTGCAACACAGAGGTTTCCAAACCTTCAAGAGTTCAATTTGTCAC 321
Db 305 TTCCATGAGCTCAGCTTCATAGACATGTTTCAATCTTCAGGAGGTACACAATGTACC 364
QY 322 CTTATGATATCAAACTGTTGTTACCGAGGACTGTAACTACTGTGCCCCAATCACAG 381
Db 365 CTCCTCTCCATAAAGACTGGTGGCTGTAGTGAAGACTGTTTATATGCTCCTCAGTCTCG 424
QY 382 CGTTACAACACTGGTGTCAAGGCTCAAGAAATCATCCAAGTTGATGAGGTGATTGAAGCT 441
Db 425 AGATATAGCATGGAGTTAAGGCAAAAGACTCATGCTTAGGAGCGTGTCAATTGATGT 484
QY 442 GCAAGAGGCAAGGCGCAATGGATCTACAAGTTCTGTATGGGTGCTGCTTGGAGAGAG 501
Db 485 GCTAAGAGGCAAAAGAGCTGGGAGCACACGTTTTTCATGGGTGCTGCTTGGCGAGAT 544
QY 502 ATGAAGGTAGAACTCAAACTTGAGAAAATCAAGAGATGATCACTGCTGTCCATGAC 561
Db 545 ACAATTGGACGGAAAAACCAACTTCAGCCAGATTTCTTGAATACATCAAGAAAATAAGAGGC 604
QY 562 CTGGAAATGGAGAGTTGTGTCACCTGGGAATGGTTGATTAAGACCAAGCCACTGAATTG 621
Db 605 ATGGGATGGAGTTTGTGTCACCTTAGGCATGATTGAGAAACAACAGCACTAGAGCTA 664
QY 622 AAAAGTGTGGTTGACGGCTACAAACCATTAACATTGATCTTCAAGGAAACACTATCCA 681
Db 665 AAGAAGGCTGCGCTCACTGCTTATAACCAATCTTGATATCTTCAAGAGAGTACTACCCA 724
QY 682 AAGTGATCTCCACAGAGCTTTGATGATAGATTGAAAACATTCAAAAACGTTCAAGGA 741
Db 725 AACGTCATCACTACTAGAAGTTATGACGATCGCTTGAACACTCTTAGCCATGTTTCTGTAT 784
QY 742 TCTGATTAAGGATGACACAGGTGTTATCTTGTGTTGGTGTGAGACCAAGAGGACCGT 801
Db 785 GCTGAATCAACGTTTGTTCAGGAGGAATCATAGGCTTGGTGGAGGAGAGAGACAGA 844
QY 802 GTATCTTTCTCTACACCTTGGCCACAAATGATCAGCATCCAGAGTCTCTTCCCAATCAAC 861
Db 845 ATAGTTTATACACAGCTGCGCAATCTTCTCTTCACTCCCTGAGAGTGTTCCTCATTAAT 904
QY 862 AGATGTTTCCCAATCAAGGCGCACGCCAATGTATGAAGAAAGTTAAGAACAAAGAGTTGAA 921
Db 905 GCTCTACTTGAGTGAAAGGTACTCTCTTGAAG-----ACCAAGAGCCAGTTGAG 955
```


AUTHORS Castelli V., Aury J.M., Jaillon O., Wincker P., Clepet C., Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V., Temple G., Caboche M., Weissenbach J. and Salanoubat M.

TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1306)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 131 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length <http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES Location/Qualifiers

1..1306

 /organism="Arabidopsis thaliana"

 /mol_type="mRNA"

 /strain="Col-0"

 /db_xref="taxon:3702"

 /clone="GSLTFB342B08"

 /tissue type="Flowers and buds"

 /plasmid="pCMVSPORT 6"

 complement(1..1306)

 /gene="At2g94360"

gene

ORIGIN

Query Match 24.8%; Score 294.4; DB 3; Length 1306;

Best Local Similarity 58.0%; Pred. No. 6e-72;

Matches 545; Conservative 0; Mismatches 386; Indels 9; Gaps 1;

QY 202 AACACCTGGACCAAGAAATAAAGCTATATATGACACACCACTCTGACCTTGATG 261

DB 219 AAGGATTGGAGTAGAGATGAATCAAGTCTGTATATGATTTCTCTCTTGACCTTC 278

QY 262 CACTATGCTCAGTGCACACAGAGGTTCCAAACACCTTCAGAGGTTCAATTGTCAC 321

DB 279 TTCCATGGAGCTCAGGTTTCATAGACATGTTTCAACTTCAGGGAGGTACAACAATGTACC 338

QY 322 CTTATGAATATCAAAACTGGTGTCTACCGAGGACTGTAAGTAACTGTGCCCAATCACAG 381

DB 339 CTCCTCTCCATAAAGACTGGTGGCTGTAGTGAAGCTGTTTCATATTTGCTCAGTCTTCG 398

QY 382 CGTTACAACACTGGTGTCAAGGCTGAAGAATCATCAAGTTGATGAGGTGATTTGAAGCT 441

DB 399 AGATAAAGCACTGGAGTGAAGGCACAAAGACTCATGCTTAAGGACGCTGTTATGATGCT 458

QY 442 GCAAGGAGGCAAGCCCAATGATCTCAAGTCTGTATGGTGTCTGCTCTTGGAGAGAG 501

DB 459 GCTAAGAAGGCAAAAGAAAGCTGGAGCACACGCTTTTTCATGGGTGCTGCTGGCGAGAT 518

QY 502 ATGAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGCTCCATGAC 561

DB 519 ACAATTGACGGAACCAACCACTTCAGCCAGATCTTTGAATACATCAAGAAATAGAGGC 578

QY 562 CTTGGAATGGAGATTGTGTCACTCCCTGGGAATGGTTGATAAAGACCAAGCCACTGAATG 621

DB 579 ATGGGGATGGAAGTGTGCTGCACCTCAGGCATGATTGAGAAACACACAGCACTAGAGCTA 638

QY 622 AAAAGTCTGGGTGACGGCGGTACACCACTATACATTGATCTTACAGGAACACTATCCA 681

639 AAGAAGGCTGGCCTCACTGCTTATAACCAACAATCTTGTATATCTTCAAGAGAGTACTACCCA 698

QY 682 AAGGTGATCTCCACAAGAAGCTTTGTATGATAGATTGAAAACATTTCAAAAACGTTTCAAGGA 741

DB 699 AAGCTCATCACTACTAGAGTATAGCAGTCGCTTGAACACTCTTAGCCATGTGCGTGAT 758

QY 742 TCTGGATTAAAGCAGTCACAGGTGTATTCTTTGGTCTTGGTGAGACCCCAAGAGGACCGT 801

DB 759 GCTGGAATCAACGTTTGTGCAGGAGCAATCATAGGGCTTGGTGAGGAGGAGAGACAGA 818

QY 802 GTATCTTCTCTACACCTTGGCCCAATGATGATCAGCATCCAGAGTCTCTTCAATCAAC 861

DB 819 ATAGGTTTATTACACAGCTGGCAACACTTCTCTCACCCTGAGAGTGTTCCTCAATTAAT 878

QY 862 AGACTGGTCCCAATCAAGGCGCAGCCCAATGTATGAAGAAGTTAAGAACAGCAAGTTGAA 921

DB 879 GCTCTACTTGCAGTGAAGGCACTCTCTTTGAAG-----ACCAGAAGCCAGTTGAG 929

QY 922 GTTGATGATGTCTGAGAACCATTTGCTACTGCAAGATTGGTCTATGCCAAGCTCTATTATC 981

DB 930 ATATGGGAGATGATCAGGATGATTGGAACCGCACGCTATTGTATGCGCAAAAGCGATG 989

QY 982 AGATTGGCTGCGAGGAGATATACAATGAAGAGGCGACAGAGTGTGCTTCATGGCT 1041

DB 990 ACACCTCTCTGCTGTAGAGTCCGGTTCTCAATGTCGGAACAAGCTCTCTGTTCTTGGCT 1049

QY 1042 GGTGTGTAATGCCATCTTCAAGGTAAAGAAATGCTCACAACATGTGTAACGCTGGAT 1101

DB 1050 GGTGCATATCTTATCTTCAACGGAGAGAGCTTTTAAACACACACATGATTTTGAC 1109

QY 1102 GAGGATAAAGCCATGTTGGCTAAATGGGCTCTGAAACCAA 1141

DB 1110 GCTGACACGCTCATGTTCAAGACATTAGGCTCATTCCTA 1149

RESULT 6

AY109523 1412 bp mRNA linear HTC 17-OCT-2002

LOCUS Zea mays CL765_1 mRNA sequence.

ACCESSION AY109523

VERSION AY109523.1 GI:21213276

KEYWORDS HTC.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 1412)

AUTHORS Hatney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes

JOURNAL Unpublished (2002)

REFERENCE 2 (bases 1 to 1412)

AUTHORS Coe, E.H.

TITLE Direct Submission

JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES Location/Qualifiers

1..1412

 /organism="Zea mays"

 /mol_type="mRNA"

 /db_xref="MaizeDB:632870"

 /db_xref="taxon:4577"

 /clone.lib="Maize Mapping Project/DuPont Cornsensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

```
ORIGIN
Query Match      24.6%; Score 292; DB 3; Length 1412;
Best Local Similarity 55.2%; Pred. No. 2.9e-71;
Matches 521; Conservative 0; Mismatches 413; Indels 9; Gaps 1;

Qy 202 AACACCTGGACCAAGAAATTAAGCTATATATGACACACCACTCATGGACTGTATG 261
    |||||
Db 242 AACGACTGGAGCGGCCGAGATCCAGCGCTACGACTACCGCTCTCTGACCTCTC 301
    |||||
Qy 262 CACTATGCTCAGGTGCAACAGAGGTTTCAACCACTTCAGAGGTTCAATTTGCACT 321
    |||||
Db 302 TTTCAGGGGCTCAGGTCCACAGAAATGTCCATAAATTCAGAGAAAGTCGCAATGCACA 361
    |||||
Qy 322 CTTATGAATATCAAACTGGTGTACCGAGGACTGTAACTACTGTGCCCCAATCACAG 381
    |||||
Db 362 CTTCTTTCAATCAAGACTGTGGATGCGTGAAGATTTGTTACTGTCTCCTCAGTCATCA 421
    |||||
Qy 382 CGTTACAACTGTGTCAAGGCTGAAAGAAATCATCAAGTTGTAGTGAAGTGAAGCT 441
    |||||
Db 422 AGATACACACTGGATTGAGGCCNNNNNTTGTGACAAAGATGCTGTNNNNNNNN 481
    |||||
Qy 442 GCAAGGAGGCAAGGCCAATGGATCTACAAGGTTCTGATGGGTGCTGTTGGAGAGAG 501
    |||||
Db 482 NNNNNNNNNNNNNNNNNCTGGAGCACCCGNNNNNNCATGGGAGCTGCGATGGAGAGAA 541
    |||||
Qy 502 ATGAAGGTAGAAAGCTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGCTCCATGAC 561
    |||||
Db 542 ACCATTGGCAGGAATCAAACTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 601
    |||||
Qy 562 CTTGAAATGGAGAGTTGTGTCACTCCCTGGGAATGGTTGATAAGACCAAGCCACTGAATTG 621
    |||||
Db 602 ATGGCATGGAGGTTCTGTGACACTAGGCATGATAGAGAAACACACAGCTGAAGAACTC 661
    |||||
Qy 622 AAAAGTGTGGTTGACGCGGTACAAACCATTAACATTGATCTTCAAGGAACACTATCCA 681
    |||||
Db 662 AAGAAGGCTGGACTTACAGCATATAATCATAACTAGATACATCAAGAGAGTATTATCCC 721
    |||||
Qy 682 AGGTGATCTCCACAGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 741
    |||||
Db 722 AACATTATATACCAAGATCATATGATGATGATGATGATGATGATGATGATGATGATGAT 781
    |||||
Qy 742 TCTGATTAAAGGCATGACAGGTGGTATTTCTTGGTCTTGGTGAGACCCCAAGAGGACCGT 801
    |||||
Db 782 GCTGGAATAGGATCTGCTCAGGTGGAATCATTTGGTCTTGGTGAAGCAGAGGAGGACCGG 841
    |||||
Qy 802 GTATCTTTCTCTACACCTTGGCCCAATGGATTCAGATCCAGAGTCTTCTTCAATCAAC 861
    |||||
Db 842 GTAGGTTGTTGCATACCTAGCTACCTTGGCTTACACACCCAGAGAGCGTTCTCTATTAT 901
    |||||
Qy 862 AGACTGTTCCCAATCAAGGCGACGCAATGTATGAAGAGTTAAGAACCAAGCAAGTTGAA 921
    |||||
Db 902 GCATTGGTTGCTGTAAGAGGACACCTCTTGGG-----ACCAAGAGCCTGTAGAG 952
    |||||
Qy 922 GTTGATGAGATTGTGAGAACCAATTTGCTACTGCAAGATTGGTTCATGCCAACGTTATTATC 981
    |||||
Db 953 ATCTGGGAATGATCCGATGATCGCACTGCTCGGATCAGATGCCAAGGCAATGTTG 1012
    |||||
Qy 982 AGATTGGCTGAGGAAGATATACATGAAGAGGAGGAGAACAGAGTGAATGTCTTCAATGGCT 1041
    |||||
Db 1013 AGGCTTTTCAAGAGCGGAGTAGTGGTTCTCGATGCCAGAACAGCGCTGTGCTTCTCGCT 1072
    |||||
Qy 1042 GGTGTAATGCCATCTTCAAGGTGAAGAAATGCTCACAACAAATGTGTAACGGCTGGGAT 1101
    |||||
Db 1073 GGGGCCAACTCCATCTTTGCGCGGAGAGAACTTCTCACAACCGCAAAACAGCACTTGTAT 1132
    |||||
Qy 1102 GAGGATAAGCCATGTTGGCTAAATGGGGTCTGTAACCAATGG 1144
    |||||
```

```
Db 1133 GCGGACGAGCGATGTTCAAGATCCTTGGCTGATCCCAAGG 1175
    |||||
RESULT 7
LOCUS      CO004857
DEFINITION EST793192 Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
            kb Coccidioides posadasii cDNA clone CIEAL82 5' end, mRNA sequence.
ACCESSION  CO004857
VERSION     CO004857.1 GI:48511746
KEYWORDS    EST.
SOURCE      Coccidioides posadasii
ORGANISM    Coccidioides posadasii
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE   1 (bases 1 to 969)
AUTHORS     Gardner, M.J. and Cole, G.T.
TITLE        Analysis of gene expression in Coccidioides posadasii mycelia and
            spherules via expressed sequence tags
JOURNAL      Unpublished (2003)
COMMENT      Contact: Gardner MJ
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301 838 3519
            Fax: 301 838 0208
            Email: Gardner@tigr.org
            Seq primer: M13 Reverse.
FEATURES     Location/Qualifiers
             1..969
                /organism="Coccidioides posadasii"
                /mol_type="mRNA"
                /strain="C735"
                /db_xref="taxon:199306"
                /clone="CIEAL82"
                /dev_stage="spherules"
                /lab_host="E. coli DH10B, T1 phage resistant"
                /clone_lib="Coccidioides posadasii spherule cDNA library,
                0.4 to 2.3 kb"
                /note="Vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;
                Coccidioides posadasii spherule cDNA library, 0.4 to 2.3
                kb"
ORIGIN
Query Match      24.5%; Score 291; DB 7; Length 969;
Best Local Similarity 60.6%; Pred. No. 5e-71;
Matches 499; Conservative 0; Mismatches 315; Indels 9; Gaps 1;

Qy 173 AATATGCAATTCTAGTGAAGACACACAGTCAACACTGGACCAAGAGAAATTAAGCTA 232
    |||||
Db 156 AAGAGCGGCTCTCAGCCACAGCACCTCGGATGAACTGGACCCGAGAGAGATTGAGAAGA 215
    |||||
Qy 233 TATATGACACACCACTCATGCACTTGATGCACTATGCTCAGGTGCACACAGAGGTTCC 292
    |||||
Db 216 TCTAGATGCACTCTGACCGAATTGACATATGCTTCGCTACGCTTACCCTGCTTCC 275
    |||||
Qy 293 AACAACTTTCAGAGGTTCAATTTGTGCACTCTTATGAATATCAAACTCGTGTGTACCG 352
    |||||
Db 276 ATGACCCCGCGCCATTCAAATGTGCACTCTCTATGACATATAAAACCGTGTGATGACGG 335
    |||||
Qy 353 AGGACTGTAACTACTGTGCCCAATCAACAGCGTTTAAACACTGGTGTCAAGGCTGAAGAA 412
    |||||
Db 336 AAGATTGCTGCTACTGCGCCCAATCTTCCCGCTACAAGACCGGGCTGAAGCCCAAAAC 395
    |||||
Qy 413 TCATCCAGTTGTAGAGTGTGATTCAGCTGCAAGAGGCAAGCCCAATGATCTACAA 472
    |||||
Db 396 TAAGCTCCGTAGACTCCGCTGCTTGAGGCGCCGAGGTCGCGAGAGAAATGGCAGCACTC 455
    |||||
Qy 473 GGTTCGTATGGGTGCTGCTTGGAGAGAGATGAAGAGTAGAAAGTCAAACTTTGAAGAAA 532
    |||||
Db 456 GATTCTGATGGGTGCGGCTTGGCGGACATGCGGAGCGGAAACCAATCTACGAACA 515
    |||||
Qy 533 TCAAGAGAGATGATCACTGCTGCTCCATGACCTTGGAAATGGAGAGTTGTGTACCCCTGGAA 592
    |||||
```

```

Db      516 TCAAGAGATGGTTTCCGGAGTCGAGCTATGGAAATGGAGTTTGGTTACTCTGGAA 575
Qy      593 TGGTTGATAAAGACCAAGCACTGAAATGAAAGTGTGGTTGACCGCGTACAAACATA 652
Db      576 TGATTGACGGCGAGCGCCAGAGGCTTAAAGAGGCTGGCTCACCGCTTACATCACA 635
Qy      653 ACATTGATCTTACAAGGAACATATCAAAAGGTGATCTCCACAAGAGCTTTGATGATA 712
Db      636 ATCTTGACAGCTCGAGAGAAATTCACCCATCCATCATCACACAGAGATCGTATGACGAGA 695
Qy      713 GATTGAAACATTCMAAAGCTTCAAGGATCTGGATTAAAGGATGCACAGGTGGTATTC 772
Db      696 GACTGCAGACCCCTGGAGCATGTCCGGATCCCGTATCAACGCTGTGTTCCGGTGGCATTC 755
Qy      773 TTGGTCTTGTGTAGACCCCAAGAGACCGTGTATCTTCCCTACACCTTGCCGCACAAATGG 832
Db      756 TGGGCTTAGGTGAACAGGACTCCGACCGCGTCCGGTTGATACACCGTCCGACATTCG 815
Qy      833 ATCAGCATCAGAGTCTCTTCCAAATCAACAGACTGGTCCCAATCAAGGGCACGCCAATGT 892
Db      816 CATCTCAACCAGAGTCTTTCCCGTGAACGCGTCTGTACCTATCAAGGGCAACCCCTG- 874
Qy      893 ATGAAGAGTTAAGAACCAAGATTTGAAGTTGATGAGATGTGCAACCAATTCGCTACTG 952
Db      875 -----GGCGAGCGGAAATGATCTCTTTTGACAAGGTTCTCCGAACCGTGGCTACGG 926
Qy      953 CAAAGATTGGTCATGCCAACTCTATTATCAGATTGGCTGCAGG 995
Db      927 CTCGTATCTTTCGCCCATGTCATTCATGTCAGATTGCCCGTGG 969

```

RESULT 8

LOCUS CO084519 876 bp mRNA linear EST 16-JUN-2004
 DEFINITION GR_Ea01D02.f GR_Ea Gossypium raimondii cDNA clone GR_Ea01D02 5', mRNA sequence.

ACCESSION CO084519

VERSION CO084519.1 GI:48775153

KEYWORDS EST.

SOURCE Gossypium raimondii

ORGANISM Gossypium raimondii

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 876)

AUTHORS Kim H., Yu Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and Wing, R. A.

TITLE Global assembly of Cotton ESTs

JOURNAL Unpublished (2004)

COMMENT Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: <http://genome.arizona.edu>

Plate: 01 row: D column: 02.

FEATURES

source

1..876
 Location/Qualifiers

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_Ea01D02"

/tissue_type="whole seedlings"

/dev_stage="first true leaves"

/lab_host="DH10B"

/clone_lib="GR_Ea"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional clones into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 24.5%; Score 290.8; DB 7; Length 876;
 Best Local Similarity 59.5%; Pred. No. 5.5e-71;
 Matches 515; Conservative 0; Mismatches 342; Indels 9; Gaps 1;
 Qy 276 GCAACACAGAGGTTTCAACAACTTTCAGAGGTTCAATTTGTGCACTCTTATGAATATCAA 335
 Db 2 GGATCACAGATATGCTCATAAATTCAGAGAGGTGCTTAGAGTACTCTCTCTCAATCAA 61
 Qy 336 AACTGCTGTTGTACCGAGGACTGTAAGTACTGTGCCCAATCACAGGTTTACAACACTGG 395
 Db 62 GACTGTGTGATGAGCAGGATTTCTTATTTGCCCTCAATCTCTAGGTACCATACTGG 121
 Qy 396 TGTCAAGGCTGAAAGAAATCATCCAACTGATGAGGTGATTGAAGCTGCAAGGAGGCAAA 455
 Db 122 CCTAAGCCCCCAACAGCTCATGACCAAGGATGCCGTTATGCAGCAGCTTAACAGGCTAA 181
 Qy 456 GGCATATGGAATCTCAAGGTTCTGTATGGGTGCTGTGGAGAGAGATGAAGGTAGAAA 515
 Db 182 AGAGGCTGTGTAGTACACGCTTTTGCATGGGTGCAGCATGGAGAGACACAGTCCGAAGAA 241
 Qy 516 GTCAAACTTTGAAGAAATCAAAGAGATGATCACTGCTGTCCATGACCTTGGAAATGAGAG 575
 Db 242 AACTAACTTCAACCAAAATCTTGAATATGTAAGAAATTAGGGATATGGGGATGGAGGT 301
 Qy 576 TTGTGTCACTCCCTGGGAATGTTGATAAAGACCAAGCCACTGAAATTCAAAAGTCTCGGTT 635
 Db 302 GTGTTCACCTTTGGGCATGCTGGAGAGCAGCAGCACTTGAATTTAAGAGGACAGTCT 361
 Qy 636 GACGGCGTCAACACCAATTAATGATACTTCAAGGAACTATCCAAAGGTGATCTCCAC 695
 Db 362 TACAGCTTACAACCAATATCTCGATACTTCAAAAGAAATATTATCCCAACATTATTACAAC 421
 Qy 696 AAGAGCTTTGATGATAGTAGTGAAGAACTTCAAAACCTTCAAGATCTGGATTAAGGC 755
 Db 422 AAGAACCCTATGATGAGCGGTTGGNAACCCCTTCAACATGTCCGTGAAGCAGGAATTAATGT 481
 Qy 756 ATGCACAGGTGATTTCTTGGTCTTTGGTGAGACCCCAAGAGGACCGGTGATCTTTCTCTA 815
 Db 482 CTGTTCAAGGGGCATTTATAGGCTTTGGAGAACAGTAGAGGACCGGTTGTTTATGCA 541
 Qy 816 CACTTGGCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 875
 Db 542 CACTTGGCTACACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCACTCCCACT 601
 Qy 876 CAAAGGACGCAACCTCTTCAAGATCAA-----AAGCCAGTGGAAATATGGGAGATGAT 935
 Db 602 GAAAGGACACCTCTTCAAGATCAA-----AAGCCAGTGGAAATATGGGAGATGAT 652
 Qy 936 CAGAACCACTGCTACTGCAAGATTGCTCATGCCAAGCTCTATTATCAGATTGGCTGCAGG 995
 Db 653 TCGGATGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 712
 Qy 996 AAGATATACAAATGAAAGAGGACAGGATGATGATGATGATGATGATGATGATGATGATGAT 1055
 Db 713 CAGAGTTCGGTCTCTATGCCAGAGGACGATTAATGTTTCTGCTGGTGAATTTCTAT 772
 Qy 1056 CTTCAAGGTAAAGAAATGCTCAACAAATGTTAAACGCTGGGATGAGGATAAAGCCAT 1115
 Db 773 CTTCACTGGTGAAGAGCTATTGACAACTCTTCAACAAATGATTATGATGCTGATCAAGTCAT 832
 Qy 1116 GTTGGCTAAATGGGGTCTGAAACCAA 1141
 Db 833 GTTCAAAATCTTGGACTGATTCCAA 858

RESULT 9

LOCUS CF822545

DEFINITION CF822545 872 bp mRNA linear EST 01-APR-2004
 EST699927 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb. Coccidioides posadasii cDNA clone CIDAG50 5' end, mRNA sequence.

```

ACCESSION      CF822545      GI:45928602
VERSION        EST
KEYWORDS      Coccidioides posadasii
SOURCE        Coccidioides posadasii
ORGANISM      Coccidioides posadasii
REFERENCE      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
AUTHORS       Onygenales; mitosporic Onygenales; Coccidioides.
TITLE         1 (bases 1 to 872)
JOURNAL       Gardner, M.J. and Cole, G.T.
COMMENT       Analysis of gene expression in Coccidioides posadasii mycelia and
              spherules via expressed sequence tags
              Other ESTs: EST699926
              Contact: Gardner MJ
              The Institute for Genomic Research
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301 838 3519
              Fax: 301 838 0208
              Email: gardner@tigr.org
              Seq primer: M13 Reverse.

FEATURES             source
    Location/Qualifiers
        1..872
            /organism="Coccidioides posadasii"
            /mol_type="mRNA"
            /strain="C735"
            /db_xref="taxon:199306"
            /clone="CIDAG50"
            /dev_stage="saprobic phase (mycelia)"
            /lab_host="E. coli DH10B, T1 phage resistant"
            /clone_lib="Coccidioides posadasii saprobic phase cDNA
            library, 2 to 4 kb"
            /note="Vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;
            Coccidioides posadasii saprobic phase cDNA library, size
            fractionated cDNA 2 to 4 kb"

ORIGIN
Query Match      22.7%; Score 270; DB 7; Length 872;
Best Local Similarity 61.3%; Pred. No. 4.5e-65;
Matches 435; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

Qy 173 AATATGCATTGTGATGAAGACACAGTCACACCTGGACCAAGAAATTAAGCTA 232
Db      |||||
Qy 163 AAGAGCGGTCTCAGCCACAGCACCTCGGATGAATGGACCGGAGAGATTGAGAAGA 222
Db      |||||
Qy 233 TATATGACACACCACTCATGTGACTTGATGCTCAGTGCAACACAGAGGTTCC 292
Db      |||||
Qy 223 TCTAGATGCCTCTGACCGAATGACATATGCTTCGCTACGTTCCCGTCTCC 282
Db      |||||
Qy 293 AACCACTTCAGAGGTTCAATTGTGCACCTTTATGAATATCAAACTCGTGGTTACCG 352
Db      |||||
Qy 283 ATGACCCCGCCGCAATTCAAATGTGCACCTCTCATGAACATAAAACCGGTGGATGCAGCG 342
Db      |||||
Qy 353 AGGACTGTAACTGTGCCCAATCACAGCGTTACACACTGGTGTCAAGGCTGAAGAA 412
Db      |||||
Qy 343 AAGATTGTCTGTACTGCGCCCAATCTCCCGCTACAAGACCGGGCTGAAGCCACAAAC 402
Db      |||||
Qy 413 TCATCCAAAGTTGATGAGGTGATTGAAGCTGCAAGAGGCAAGGCCAATGGATCTCAA 472
Db      |||||
Qy 403 TAAGCTCCGTAGACTCCGTGCTTAGGCGCGCAGGGTGCACAGAGAAATGGCAGCATC 462
Db      |||||
Qy 473 GGTTCGTATGGGTGCTGCTTGGAGAGAGATGAAGAGGTAGAAGATCAAACTTGAAGAAAA 532
Db      |||||
Qy 463 GATTCTGATGGTGGCCGCTTGGCGGACATGCGAGGCGGAAACCAATCTACGAACA 522
Db      |||||
Qy 533 TCAAAGAGATGATCATCTGCTTCATGACCTTGGAAATGGAGAGTTGTGTCAACCTGGGAA 592
Db      |||||
Qy 523 TCAAAGAGATGGTTTCGGAGTGCAGACTATGGGAATGGAATTTGCGTTACTCTGGGA 582
Db      |||||
Qy 593 TGGTTGTAAGAACCAAGCCACTGAATTGAAGAGTGTGGGTGAGCGGTACACACCATTA 652
Db      |||||
Qy 583 TGATTGACGGGAGCGGCAAGAGCTTAAAGACGCTTGGCTTACCGCTTACATATCA 642
Db      |||||
Qy 653 ACATTGATATTTACAAGGAACACTATCCAAAGGTTGATCTCCACAGAGCTTTTGATGATA 712
Db      |||||

Db 643 ATCTTGACAGCTCGAGAGAATTTACCCATCATCACCCAGAGATCGTATGACGAGA 702
Qy 713 GATTGAACAATTCAAAACGTTTCAAGGATCTGATTAAAGGCATGACACAGTGTGTTTC 772
Db      |||||
Qy 703 GACTGCAGACCTTGGAGCATGTCCGCGATGCGGTATACAGCTGTTCGGGTGGCATTC 762
Db      |||||
Qy 773 TTGGTCTTGGTGAGACCCAAAGAGACGCTGTATCTTTCTCTACACCTTGGCCCAATGG 832
Db      |||||
Qy 763 TGGGCTTAGGTGAACAGGACTCCGACCGGTGCGGTTGATACACACCGTCCACATTCG 822
Db      |||||
Qy 833 ATCAGATTCAGATCTCTTCCCAATCAACAGACTGGTCCCAATCAAGGGC 882
Db      |||||
Qy 823 CATCTCACCCAGAGTCTTTCCCGTGAAACGCGCTCGTACCTATCAAGGGC 872
Db      |||||

RESULT 10
CF822544/c
LOCUS
DEFINITION EST699926 Coccidioides posadasii saprobic phase cDNA library, 2 to
4 kb Coccidioides posadasii cDNA clone CIDAG50 3' end, mRNA
sequence.
ACCESSION      CF822544      GI:45928601
VERSION        CF822544
KEYWORDS      EST.
SOURCE        Coccidioides posadasii
ORGANISM      Coccidioides posadasii
REFERENCE      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
AUTHORS       Onygenales; mitosporic Onygenales; Coccidioides.
TITLE         1 (bases 1 to 834)
JOURNAL       Gardner, M.J. and Cole, G.T.
COMMENT       Analysis of gene expression in Coccidioides posadasii mycelia and
              spherules via expressed sequence tags
              Unpublished (2003)
              Other ESTs: EST699927
              Contact: Gardner MJ
              The Institute for Genomic Research
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301 838 3519
              Fax: 301 838 0208
              Email: gardner@tigr.org
              Location/Qualifiers
                  1..834
                      /organism="Coccidioides posadasii"
                      /mol_type="mRNA"
                      /strain="C735"
                      /db_xref="taxon:199306"
                      /clone="CIDAG50"
                      /dev_stage="saprobic phase (mycelia)"
                      /lab_host="E. coli DH10B, T1 phage resistant"
                      /clone_lib="Coccidioides posadasii saprobic phase cDNA
                      library, 2 to 4 kb"
                      /note="vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;
                      Coccidioides posadasii saprobic phase cDNA library, size
                      fractionated cDNA 2 to 4 kb"

ORIGIN
Query Match      22.4%; Score 266; DB 7; Length 834;
Best Local Similarity 61.6%; Pred. No. 5.1e-64;
Matches 447; Conservative 0; Mismatches 270; Indels 9; Gaps 1;

Qy 430 GTGATTGAAGTGCAAAAGGAGGCAAGCCAAATGGATCTACAGGTTCTGTATCGGTGCT 489
Db      |||||
Qy 827 GTGCTTGAGCGCGCCAGGTCGCCAAGAGAAATGGCAGCACTCGATTTCTGATCGGTGCC 768
Db      |||||
Qy 490 GCTTGGAGAGAGATGAAGAGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACT 549
Db      |||||
Qy 767 GCCTGGCGGACATGCGAGGCGGAAAAACCAATCTACGAAACATCAAGAGATGTTTCC 708
Db      |||||
Qy 550 GCTGTCCATGACCTTGAATGGAGAGTTGTGTCAACCTCGGGAATGGTTGATAAGACCAA 609
Db      |||||
Qy 707 GGAGTGCAGCTATGGGAATGGAAAGTTTGCTTACTCTGGGAATGATGACGGCGAGCAG 648
Db      |||||

```

```

QY 610 GCCACTGAATTGAAAGTCTGGTTCACGGCTGACCAACCATTAACATTTGATCTTACAG 669
Db 647 GCCAGAGCTTAAGACGCTGGCCCTCACCGCTTACAATCACAATCTTGACAGTCGAGA 588
QY 670 GAACACTATCCAAAGGTGATCTCCACAAGAGCTTTGATGATAGATTGAAAACATTCAAA 729
Db 587 GAATTTACCCATCCATCATCACACAGAGATCGTATGACGAGAGACTGCAGACCTTGGAG 528
QY 730 AAGCTTCAAGGATCTGGATTAAAGGATGACAGGTGGTATTTCTTGGTCTTTGGTGAGACC 789
Db 527 CATGTCGGCGATGCCGATCAACGTCGTTCGGGTGGCATTTCTGGGCTTAGGTGAACAG 468
QY 790 CAAGAGACCGGTATCTTCTCTACACTTGGCCACATGATGATGATGATGATGATGATGAT 849
Db 467 GACTCGGACCGGTCTGGGTGATACACACCGTCCGACATTTGCCATCTCACCAGAGTCT 408
QY 850 CTTTCAATCAACAGACTGGTCCCAATCAAGGGCAGCCCAATGATCAAGAAGTTAAGAAC 909
Db 407 TTCCCGGTGAAACGGCTCGTACTATCAAGGGCAGCCCTGGCGAGCGGAAATGATC 348
QY 910 AAGCAAGTTGAAGTTGATGAGATTGTGAGAACCATTTGCTACTGCAAGATTGGTCAATGCCA 969
Db 347 TCTT-----TTGACAAGGTTCTCCGAAACGGTGGCTACGGCTCGTATCGTTCTGCC 297
QY 970 AGCTCTATTATCAGATTGGCTGCGAGAGATATACATGAAGAGGCGACAGGATGATG 1029
Db 296 ATGTCCATTGTGAGATTGGCTGCGAGAGATATTCGATAGCTGGGATATTAACCCATGGCACCT 237
QY 1030 TGTCTCATGGCTGGTGTAAATGCTATCTTCAAGGTAAAGAAATGCTCAACAATGTGT 1089
Db 236 TGTCTTCCAGGCTGGTGGATGGATCTTCAAGGAGAGAGATGTTAATCCAGACTGC 177
QY 1090 AACGGCTGGATGAGGATAAAGCCATTTGGCTTAAATGGGCTGTAACCAATGGAGAT 1149
Db 176 TCCGGTGGATGAGGATAAAGATATTCGATAGCTGGGATATTAACCCATGGCACCT 117
QY 1150 TTCAA 1155
Db 116 TTTGAA 111

```

RESULT 11

```

CF721966/c
LOCUS
DEFINITION
CF721966 C.neoformans strain JEC21 Cryptococcus neoformans var.
neoformans cDNA clone CCA1517, mRNA sequence.
CF721966
VERSION
CF721966.1 GI:41576125
KEYWORDS
SOURCE
Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
var. neoformans)

```

ORGANISM

```

Cryptococcus neoformans var. neoformans
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
Filobasidiella.

```

REFERENCE

```

1 Loftus, B.
End sequencing of clones from a Full length enriched, normalized
JEC21 cDNA library
Unpublished (2003)
Other ESTs: CCA1517R
Contact: Brendan Loftus
TIGR

```

JOURNAL

```

COMMENT
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: cryptotigr.org
Seq primer: TF.
Location/Qualifiers
1..829
/organism="Cryptococcus neoformans var. neoformans"
/mol_type="mRNA"

```

FEATURES

```

source
CF719376
CF719376/c
LOCUS
DEFINITION
CF719376 C.neoformans strain JEC21 Cryptococcus neoformans var.
neoformans cDNA clone CCAEF16, mRNA sequence.
CF719376
ACCESSION

```

ORIGIN

```

Query Match      22.1%; Score 262.4; DB 7; Length 829;
Best Local Similarity 60.6%; Pred. No. 6.4e-63;
Matches 453; Conservative 0; Mismatches 286; Indels 9; Gaps 1;

QY 413 TCATCCAAAGTTGATGAGGTGATTGAAGCTGCACAAAGGAGGCAAGGCAATGATCTACAA 472
Db 821 TCGTCAGCATTTGAGCGCTGTGCTTGAAGCTGCTCGTCAGCCAGGCCACGCTCCACTA 762
QY 473 GGTTCCTGATGGGTGCTGCTTGGAGAGAGATGAAAAGGTAGAAAGTCAAACTTTGAAGAAA 532
Db 761 GATTCCTGATGGGTGCTGCTTGGAGAGACTTTGGCCGCTAGAAAGAGTGGATTTCGAAAAG 702
QY 533 TCAGAGAGATGATCACTGCTGCTCCATGACCTTGGATGGAGAGTGTGTCAACCTGGGAA 592
Db 701 TCTTGAATGGTTAAAGAGGT/CAGGGGTATGGGAATGGAGGTCTGCACGACTTTGGGTA 642
QY 593 TGGTTGATAAAGACCACGACCTGAATTTGAAAAGTGCTGGGTGACGGGTGACCAACCAT 652
Db 641 TGCTCTCTCTGAGCAGCTATCCGGCTCAAGGAAGCTGGTCTCAGCGCTTATACCA 582
QY 653 ACATTGATATCTTCAAGGAAACATATTCCTCAAGAGTGTATCCCAAGAGAGCTTTGATGATA 712
Db 581 ACCTTGACACTTCTCGAATTTCTATCCGAGGTGTCTCCTCAGATCATACGATGATC 522
QY 713 GATTGAAAACATTCAAACCGTTCAAGGATCTGGATTAAAGGATGACAGGTGGTATTC 772
Db 521 GATTGACCACTTGTCTGCCGTGGAGAGCTGGTATCTCAGTCTGTCTTGGTGGCATCC 462
QY 773 TTGTTCTTGGTGAGACCCAGAGGACCGTGTATCTTCTCTACACCTTGGGCAACATGG 832
Db 461 TCGGTTTGGGAGAGGGGATGAGATCGTTCGGATTGATCCACGAGGTTTCCAGGATGC 402
QY 833 ATCAGATCCAGAGTCTTCTTCAATCAACAGACTGGTCCCAATCAAGGGCAGCCCAATGT 892
Db 401 CCGAGCACCTTGAGTCTTTCCTCCGCTCAACACATGTCTCCTATCCTGCTACTCCCTCG 342
QY 893 ATGAAGAGATTAAAGAACAGCAAGTTGAAGTTGATGAGATTGTTCAGAACCATTTGCTACTG 952
Db 341 AAGGC-----AAGAGCCTGTCAAGGTTCCACACCGCTCTCCGAACCATCGCTACTG 291
QY 953 CAAGATTGGTTCATGCCAACGTCATTATTCAGATTGGCTGAGGAGATATACAATGAAAG 1012
Db 290 CCGTATCGTCTTCCCAAGACCATCATTCGTCTCGCGCTGGCGGACACGAATTCAGCG 231
QY 1013 AGGCAGAACAGGTGATGTGCTTCAATGAGTGGTGGTGAATGCCATCTTCAAGGTGAAGAAA 1072
Db 230 AGACCGAACAGCCATGSCCTTTCATGSCGCTGCCAATGCCATCTTCACTGCGAGAGA 171
QY 1073 TGCTCACACATGTGTAAACGGCTGGGATGAGGATGAAGCCATGTGGCTAAATGGGGTC 1132
Db 170 TGCTCACACACCCCTTGTCTCCGGATGGGATGAGGACAAAGGCTATGTGTGACAGGTGGGTC 111
QY 1133 TGAACCAATGGAGAGTTTCAAAATACAA 1160
Db 110 TCAGGGCCAGAGAGAGCTTTGAGGACAA 83

```

RESULT 12

```

CF719376/c
LOCUS
DEFINITION
CF719376 C.neoformans strain JEC21 Cryptococcus neoformans var.
neoformans cDNA clone CCAEF16, mRNA sequence.
CF719376
ACCESSION

```

```

VERSION      CF719376.1  GI:41573535
KEYWORDS     EST.
SOURCE       Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
              var. neoformans)
ORGANISM      Cryptococcus neoformans var. neoformans
              Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
              Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
              Filobasidiella.
REFERENCE     1 (bases 1 to 862)
AUTHORS       Loftus,B.
TITLE         End sequencing of clones from a full length enriched, normalized
              JEC21 cDNA library
JOURNAL       Unpublished (2003)
COMMENT       Other ESTs: CCAEP16TR
              Contact: Brendan Loftus
              TIGR
              Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
              Rockville, MD 20850, USA
              Tel: 301-838-3543
              Fax: 301-838-0208
              Email: cryptotigr.org
Seq primer: 1F.
              Location/Qualifiers
FEATURES             1..862
              /organism="Cryptococcus neoformans var. neoformans"
              /mol_type="mRNA"
              /strain="JEC21"
              /db_xref="taxon:40410"
              /clone_lib="C.neoformans strain JEC21"
              /note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full
              length, normalized library was prepared from a variety of
              conditions using RNA provided by Joseph Heitman and
              Jennifer Lodge"
ORIGIN
Query Match      21.2%; Score 252.4; DB 7; Length 862;
Best Local Similarity 60.5%; Pred. No. 4.5e-60;
Matches 437; Conservative 0; Mismatches 276; Indels 9; Gaps 1;

Qy 439  GCTGAAAGAGGCAAGCCAAATGGATCTCAAGGTTCTGTATGGGTGCTGCTTGGAGA 498
      |||||
Db 862  GCTGCTCGTCAAGCCAAAGCCAAACGGCTCCATAGATTCTGATGGGTGCTGCTTGGAGA 803
      |||||
Qy 499  GAGATGAAGGTAGAAAGTCAAACTTGAAGAAATCAAGAGATGATCACTGCTGTCCAT 558
      |||||
Db 802  GACTTGGCCGCTAGAAAGAGTGGATTTCGAAAGATCTTTGAAATGTTTAAAGAGGTGAG 743
      |||||
Qy 559  GACCTTGAATGGAGAGTGTGTCAACCTGGGAATGGTTGATAAGACCAAGCCACTGAA 618
      |||||
Db 742  GGTATGGAAATGGAGGCTCTGCACGACTTTGGGTATGCTCTCTCTGAGCAAGCTATCCGG 683
      |||||
Qy 619  TTGAAAAGTGTGGGTGACGGGCTPACCAACCATTAACATTGATTAACAAGGAACACTAT 678
      |||||
Db 682  CTCAAGGAAGCTGGTCTCAGCGCCTATAAACCAACCTTGACACTTCTCGAGAATTCCTAT 623
      |||||
Qy 679  CCAAGGTGATCTCCAGAGAGCTTTGATGATGATTAAGAAATTCATAAAAGGTTCAA 738
      |||||
Db 622  CCCGAGGTGTGCACCTCAAGATCAACGATGATGATGATGATGATGATGATGATGATGAT 563
      |||||
Qy 739  GGATCTGGATTAAGGCAATGACAGGTGGTATCTTGGTCTTGGTGAACCAAGAGGAC 798
      |||||
Db 562  GAAGCTGGTATCTAGTCTGTCTTGGTGGCATCTCGGTTTGGGAGCGGGATGAGGAT 503
      |||||
Qy 799  CGTGTATCTTCTCTACACCTTGGCCACATGAGATCCAGATCTCTTCTTCCAAATC 858
      |||||
Db 502  CGTCTGGATGATTCACAGGTTTCCAGGATGCCGAGCACCTTGAGTCTTCTCCCGTC 443
      |||||
Qy 859  AACAGATGCTGCCAATCAAGGCGACCGCATGATGATGAAGATTTAAGAACCAAGATT 918
      |||||
Db 442  AACACATTTGCCCTATCCCTGGTACTCCCTTCGAAGGC-----AACGCGCTGTC 392
      |||||
Qy 919  GAAGTTGATGAGATTGTGCAAGCAACATTGCTACTGCAAGATTGGTCAATGCCAACGTCATT 978

```

```

Db 391  AAGGTCCACACCGTCCTCCGAACCATCGCTACTGCCGATATGTTCTTCCCAAGACCATC 332
      |||||
Qy 979  ATCAGATTGCTGCAGGAAGATATACATGAAGAGCAGACAGGTGATGCTTTCATG 1038
      |||||
Db 331  ATTGCTTCGCGCGTGGCCGACACGAATTCAGCAGACCGAACCAAGCCATGGCTTTCATG 272
      |||||
Qy 1039  GCTGTTGTAATGCAATCTTTCACAGGTAAGAAAAATGCTCAACAATGTGTAACGGCTGG 1098
      |||||
Db 271  GCGGTCCCAATGCCATCTTCACTGGCGAGAGATGCTCACCACCCCTTGTCCGGATGG 212
      |||||
Qy 1099  GATGAGGATTAAGCCATGTTGGCTTAATGGGCTGTAACCAACCAATGGAGAGTTCAATATAC 1158
      |||||
Db 211  GATGAGGACAAGGCTATGTTGGACAGGTGGGCTCTCAGGGGCCAGAGAAAGCTTTGAGGAC 152
      |||||
Qy 1159  AA 1160
      ||
Db 151  AA 150

RESULT 13
CD852378
LOCUS       DH0ALL272F09Z2M1 HaDevR5 Helianthus annuus cDNA clone HaDevR527F09,
DEFINITION mRNA sequence.
ACCESSION   CD852378
VERSION     CD852378.1  GI:32536194
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
ORGANISM    Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
REFERENCE   1 (bases 1 to 781)
            Genopiante.
            Genopiante, a major partnership french program in plant genomes
            Unpublished (2003)
            Contact: Genopiante
            Genopiante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genopiante' (http://www.genopiante.com
            and http://genopiante-info.infobiogen.fr).
FEATURES             1..781
              Location/Qualifiers
              /organism="Helianthus annuus"
              /mol_type="mRNA"
              /culturvar="psc8"
              /db_xref="taxon:4232"
              /clone="HaDevR527F09"
              /tissue_type="4 days after self-pollination embryo"
              /clone_lib="HaDevR5"
ORIGIN
Query Match      21.2%; Score 251.6; DB 6; Length 781;
Best Local Similarity 58.5%; Pred. No. 7.3e-60;
Matches 462; Conservative 0; Mismatches 319; Indels 9; Gaps 1;

Qy 209  GGACCAAGAGAAATTAAGCTATATATACACACCACCTCATGAGCTTGTATGCACATG 268
      |||||
Db 1  GGAAACAAGAGAGATCAAGTCCGTTTATGATTTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
      |||||
Qy 269  CTCAGGTGCAACAGAGAGGTTCACACACCTTCAGAGGTTCAATTTGCACTCTTATGA 328
      |||||
Db 61  GAGCTCAAGTTCTATAGACATGCTCAGAAATTTAGGAGGTTTCAGCAGTGTACGCTTCTGT 120
      |||||
Qy 329  ATATCAAACTGGTGGTTGTACCGAGGACTGTAAGTACTGTGCCCATCACAGCGTTACA 388
      |||||
Db 121  CTATAAAGACGGGTGGGTGCAGTGAAGATTGTTCTTATGCCCTCAATCTTCTAGGTATG 180
      |||||

```

Qy	389	ACACTGGTGTCAAGGCTGAAAGAAATCATCCAAAGTTGATGAGGTGATGTAAGCTGCACAAAGG	448
Db	181	ATACTGGAGTGAAGAGCTCAAAAGCTCATGAACAAGGATCTGTTCTTGAAGCTGCACAAA	240
Qy	449	AGGCAAAAGCCCAATGGATCTACAAGGTTCTGTATGGTGCTCTCTGGAGAGAGATCAAG	508
Db	241	AGGCCAAAAGAGCTGGCAGCATCTCGGTTTTGCATGGTGCTCGGTGGAGAGACACAATAG	300
Qy	509	GTAGAAAGTCAAACTTGGAAGAAAATCAAGAGATGATCACTGCTGTCCATGACCTTGGAA	568
Db	301	GAAGGAAAACCAACTTTAGTCAGATCCTTGAATATGTGAAGAGATAAGGGGTATGGGAA	360
Qy	569	TGGAGAGTTGTGTCAACCTGGGAATGGTTGATAAAGACCAAGCCACTGAAATTGAAAAGTG	628
Db	361	TGGAGGTTTGTGTACGTGGGAATGATAGAAAAGCAGCAGGCTTTAGAACTCAAGAAG	420
Qy	629	CTGGGTTGACGGCGTACAACCATTAACATTGTACTTTCAAGGAACACTATCCAAAGGTGA	688
Db	421	CAGGCTTACAGCTTACAAATCATAACTTGTATACATCAAGAGAGTACTACCTTAATCA	480
Qy	689	TCTCCAAGAAGCTTTGATGATAGATTGAAAACAATCAAAAACGTTCAAGGATCTGGAT	748
Db	481	TTACAACAAGAAACATATACGAAACGCTTGGAAACCATCAAGCATGTTGTTGAAGCAGGAA	540
Qy	749	TAAAGGCATGCACAGGTGGTATTCTTGGTCTGGTGAGACCAAGAGGACCGTGTATCTT	808
Db	541	TTAATGTTTGTCTGGAGGAATAATAGGGCTTGGTGAAGACAGAGACAGGGTTGGGT	600
Qy	809	TCCTCTACACCTTGGCCACAATGGATCAGCATCCAGAGTCTCTTCCAATCAACACACTGG	868
Db	601	TGTTGCTATACATTAGCAACACTCCCTCACACCCGGAGAGTGTCCCATTAACGCACTTC	660
Qy	869	TCCCAATCAAGGGCAGCCCAATGTATGAAGAAGTTAAGAAACAAGCAAGTTGAAGTTGATG	928
Db	661	TTGCAGTTAAAGGCACACCTTTTGGAAAGTCAA-----AAGGCAGTTGAGATATGGG	711
Qy	929	AGATTGTGAGAACCAATTCCTACTGCAAGATTGGTTCATGCCAACGTCATTATTCAGATTGG	988
Db	712	AGATGATACCGATGATTTGCTACGGCCCGCATACGATGCCAAAGCTATGGTGAGGTTAT	771
Qy	989	CTGCAGGAAG	998
Db	772	CTGCAGGAG	781

BU063672	770 bp	mrna	linear	EST 26-AUG-2002
LOCUS	Fgr_3 G19_T7	Nitrogen-starved mycelia	Gibberella zeae cDNA, mRNA	
DEFINITION	sequence.			
ACCESSION	BU063672			
VERSION	BU063672.1			
KEYWORDS	EST.			
SOURCE	Gibberella zeae			
ORGANISM	Gibberella zeae			
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.			
AUTHORS	1 (bases 1 to 770)			
TITLE	Trail,F., Xu,J.-R., San Miguel,P., Halgren,R.G. and Kistler,H.C. Analysis of expressed sequence tags from Gibberella zeae (anamorph <i>Fusarium graminearum</i>)			
JOURNAL	Fungal Genet. Biol. 38 (2), 187-197 (2003)			

FEATURES

source
1. 770
/organism="Gibberella zeae"
/mol_type="mRNA"
/strain="NRRL 31084"
/db_xref="taxon:5518"
/clone_lib="Nitrogen-starved mycelia"
/notes="vector: Uni-ZapII; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 21.2% Score 251.4; DB 5; Length 770;
Best Local Similarity 62.7%; Pred. No. 8.3e-60;
Matches 412; Conservative 0; Mismatches 236; Indels 9; Gaps 1;

QY 495 GAGAGAGATGAAAGGTAGAAAGTCAAACCTTTGAAGAAAAATCAAAGAGATGATCACTCTCTG 554
|||||
Db 756 GAGGACATGCGCGGACGAAGAACGCCCTCAAGAACATCAGGCCCATGTCGAGGGCGT 697
|||||
QY 555 CCATGACCTTGGATGGAGAGTTGTGTACCCCTGGGAATGTTGATAAAGACCAAGCCAC 614
|||||
Db 696 CAAGGGCATGGGAATGGAGGTCCTGCGTAAACGCTGGGCATGTCGACGCGAGCAGGCCAA 637
|||||
QY 615 TGAATCGAAAGTCTGGGGTTGACGGCGTACACCAATCAATGATGATCTTACAGGAACA 674
|||||
Db 636 GGAGCTCAAGCGCGCAGGACTCAGCGCGTCAACCAACAATGTTGATCTAGCCGAGGTT 577
|||||
QY 675 CTATCCAAAGGTGATCTCCACAGAAAGCTTTGATGATAGATTGAAAAAATTCAAAAACGT 734
|||||
Db 576 TTACCCCAACGTCATCAGACGGCGAGCTACGACGAGCGACTGCAGACACTGATGTCATGT 517
|||||
QY 735 TCAAGGATCTGGATTAAGGCAATGCAAGGTGGTATCTTTGGTCTTGGTGAGACCCAGA 794
|||||
Db 516 CCGTGATGCTGGAAATTAATGTCTGCTCTGGTGAATTCGGGTCTTGGTGAGACGACGA 457
|||||
QY 795 GACCGGTGATCTTCTCTACACCTTGGCCCAATGGATCAGCATCCAGAGTCTCTTCC 854
|||||
Db 456 GACCGGTGTTGGTTTGTCTCATACCGTCTCGACGCTCCCAAGCCACCCGAGAGCTTCCC 397
|||||
QY 855 AATCAACACAGACTGGTCCCAATCAAGGCGACGCCAATGTATGAAGAAGTTAAGAACCAAGCA 914
|||||
Db 396 CGTGAATGCCCTGGTCTCTATCAAGGTTACACCACTTGGTGACACCGATG----- 345
|||||
QY 915 AGTTGAAGTTGATGAGATGTCAGAACCAATGCTACTGCAAGATTTGGTCATGCCAACGTC 974
|||||
Db 344 -GTGGAGTTTACGAGCATGCTCGGACCATCGCGACCGCGGCATCATCATCGCTCGAC 286
|||||
QY 975 TATTATCAGATTGGCTGACGAGGATATACATGAAGAGGCGAGACAGGTCATGTGCTT 1034
|||||
Db 285 CATCATCCGATCGCAGCGGGTTCGAAGACCATGTCTAGAGGAGAGCAAGCGCTATGCTT 226
|||||
QY 1035 CATGGCTGGTTGATGTCATCTTCCACAGTAAAGAAATGCTTCACAACAATGTGTAAACGG 1094
|||||
Db 225 CATGGCAGTGCACACGCCATCTTTACGGGTGAGAAGATGCTGACGACCGAGTGTATGG 166
|||||
QY 1095 CTGGGATGAGGATAAAGCCATGTTGGCTAAATGGGCTCTGAACCAATGGAGAGTTT 1151
|||||
Db 165 CTGGGATGAGGATGGGCCATGTTTGGTCATGGGCTCTGGAGGCCATGAAGAGCTT 109
|||||

RESULT 15
CA781203
LOCUS
DEFINITION
CA781203 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA
sequence.
CA781203
CA781203.1 GI:26019227
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 671)
Lundegaard,M.,Emmersen,J.,Nielsen,K.L.,Wilson,I.,Somerville,S.

TITLE and Welinder, K.G.
EST sequencing of Erysiphe cichoracearum infected Arabidopsis
plants
JOURNAL Unpublished (2002)
COMMENT Contact: Karen G. Welinder
Institut for Biotechnologi
Aalborg Universitet
Sohnsgaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.au.dk.

FEATURES
source
1. .671
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/dev_stage="Plant 3 weeks old, three days post infection"
/clone_lib="Infected Arabidopsis Leaf"
/note="Organ: Leaf; Vector: pBluescript; Mixed cDNA
library of Arabidopsis and E. cichoracearum infected leaf
from three weeks old Arabidopsis plants. Plants were
harvested 3 days after infection and mRNA oligo dT
selected."

ORIGIN
Query Match 21.1%; Score 250.2; DB 6; Length 671;
Best Local Similarity 64.0%; Pred. No. 1.7e-59; Indels 0; Gaps 0;
Matches 378; Conservative 0; Mismatches 213;
QY 306 GGTTCAAATTTGTCACCTCTTATGAATATCAAACTGGTGGTTGTACCGAGGACTGTAAAGTA 365
Db 31 GATCAGATGTGTACACTCATGAACATCAAACTGGTGGCTGTAGCGAAGATTGCTCTTA 90
QY 366 CTGTGCCCAATCACAGCTTACACACTGGTGTCAAGGTGAAAGAAATCATCCAGTTGA 425
Db 91 TTGTGCTCAATCTTCTCGCTATGATACAGGATTTAAAGGCCCAACAGATGGTAACTGTTGA 150
QY 426 TGAGGTGATTGAAGCTGCAAGGAGGCAAGGCCAATGGATCTACAAAGTTCTGTATGGG 485
Db 151 CTCTGTCTTAGAAGCAGCCAAATTCGGAAGAGTAACGGTAGCACAAAGATTCTGTATGGG 210
QY 486 TGCTGCTTGGAGAGATGAAGGTAGAAAGTCAAACTTCAAGAAATCAAAAGAGATGAT 545
Db 211 TCGGCTTCGAGAGACATCGAGGACGAAGTCGGCTTGAAGAACATCAAGGAATGGT 270
QY 546 CACTGTGTCATGACCTTGGAAATGAGAGATTGTGTACCCCTGGGAAATGGTTGATAAGA 605
Db 271 TAGTGGGGTTCGAAGTTTGGGTATGGAATGTTGTGACGCTAGGAATGATCGACCTTGA 330
QY 606 CCAAGCCACTGAATTCAAAAGTCCTGGGTTGACGGGTACACCATAACATTGATACTTA 665
Db 331 GCAGGCAAAAGAAATTAAGAGCGAGGCTTGACAGCATATATCATATGTCGATACATC 390
QY 666 CAAGGAACACTATCCAAAGGTGATCTCCACAAGAGCTTTGATGATAGATTGAAAACATT 725
Db 391 AAGGGAACACTACCCCTCAGTAATACTACTCGAACCTACGATGAGCGACTTCAAACTAT 450
QY 726 CAAAACAGTTCAGGATCTGGATTAAAGGATGCACAGGTGGTATTTCTTGGTCTTGGTGA 785
Db 451 AGCAAAACGTGCGGGATGCGAGGCTTAAACGTTTGTACTGGAGGCAATTTCTAGGCTGGGTGA 510
QY 786 GACCCAGAGAGGACCGGTATCTTTCCTCTACACCTTGGCCACAATGGATCAGATCCAGA 845
Db 511 AAAGCAGAGGACCAATTTGGCCCTAATATACACTGTCCCACTATGCCAGCACATCTCGA 570
QY 846 GTCTCTTCCAATCAACAGACTGGTCCCAATCAAGGCGAGCCCAATGTATGA 896
Db 571 GAGTTTCTCTGTTAACACGTTAGTACCACCAAGGTACACCCGCTAGGTGA 621